Joint Source Based Brain Imaging Analysis for Classification of Individuals

by

Mahdi Ramezani

B.Sc., Sharif University of Technology, 2007
M.Sc., Sharif University of Technology, 2010

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
in
The Faculty of Graduate and Postdoctoral Studies
(Electrical and Computer Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)
December 2014
© Mahdi Ramezani 2014
Abstract

Diagnosis and clinical management of neurological disorders that affect brain structure, function and networks would benefit substantially from the development of techniques that combine multi-modal and/or multi-task information. Here, we propose a joint Source Based Analysis (jSBA) framework to identify common information across structural and functional contrasts in data from MRI and fMRI experiments, for classification of individuals with neurological and psychiatric disorders. The framework consists of three components: 1) individual’s feature generation, 2) joint group analysis, and 3) classification of individuals based on the group’s generated features. In the proposed framework, information from brain neuroimaging datasets is reduced to a feature that is a lower-dimensional representation of a selected brain structure or task-related activation pattern. For each individual, features are used within a joint analysis method to generate basis brain activation sources and their corresponding modulation profiles. Modulation profiles are used to classify individuals into different categories. We perform two experiments to demonstrate the potential of the proposed framework to classify groups of subjects based on structural and functional brain data. In the fMRI analysis, functional contrast images derived from a study of auditory and speech perception of 16 young and 16 older adults are used for classification of individuals. First, we investigate the effect of using multi-task fMRI data to improve the classification accuracy. Then, we propose a novel joint Sparse Representation Analysis (jSRA) to identify common information across different functional contrasts in data. We further assess the reliability of jSRA, and visualize the brain patterns obtained from such analysis. In the sMRI analysis, features representing position, orientation and size (i.e. pose), shape, and local tissue composition of brain are used to classify 19 depressed and 26 healthy individuals. First, we incorporate pose and shape measures of morphology, which are not usually analyzed in neuromorphometric studies, to measure structural changes. Then, we combine brain tissue composition and morphometry using the proposed jSBA framework. In a cross-validation leave-one-out experiment, we show that we can classify the subjects with an accuracy of 67% solely based on the infor-
Abstract

Information gathered from the joint analysis of features obtained from multiple brain structures.
Preface

This thesis is primarily based on five journal and five conference papers, resulting from the collaboration between multiple researchers. In all publications, the contribution of the author was in developing, implementing, and evaluating the method. All co-authors contributed to the editing of the manuscript. Ethics approval for conducting this research has been provided by the Clinical Research Ethics Board, certificate numbers: PSYC-066-07, PSYC-071-07

A study described in Chapter 3 has been published in:


The contribution of the author was in developing, implementing, and evaluating the method. K. Marble, H. Trang and Dr. Johnsrude collected the dataset. Profs. Abolmaesumi and Johnsrude helped with their valuable suggestions in improving the methodology.

A study described in Chapter 4 has been published in:


- Mahdi Ramezani, Purang Abolmaesumi, Kristopher Marble, Heather Trang, and Ingrid Johnsrude, Classification of individuals based on Sparse Representation of brain cognitive patterns: A functional MRI

The contribution of the author was in developing, implementing, and evaluating the method. K. Marble, H. Trang and Dr. Johnsrude collected the dataset. Profs. Abolmaesumi and Johnsrude helped with their valuable suggestions in improving the methodology.

A study described in Chapter 5 has been submitted:


The contribution of the author was in developing, implementing, and evaluating the method. S. Noranian, Dr. Abolmaesumi and Dr. Johnsrude helped with their valuable suggestions in improving the methodology.

A study described in Chapter 6 has been published in:


The contribution of the author was in developing, implementing, and evaluating the method. Dr. Rasoulian helped with the developing of the statistical model. R. Bosma, R. Tong, Dr. Hollenstein and Dr. Harkness collected the MRI dataset. Dr. Abolmaesumi and Dr. Johnsrude helped with their valuable suggestions in improving the methodology.

A study described in Chapter 7 has been published in:

- Mahdi Ramezani, Purang Abolmaesumi, Amir Tahmasebi, Rachael Bosma, Ryan Tong, Tom Hollenstein, Kate Harkness, and Ingrid Johnsrude, Fusion Analysis of First Episode Depression: Where Brain Shape Deformations Meet Local Composition of Tissue, Neuroimage: Clinical.

The contribution of the author was in developing, implementing, and evaluating the method. R. Bosma, R. Tong, Dr. Hollenstein and Dr. Harkness
Preface

collected the MRI dataset. Dr. Abolmaesumi and Dr. Johnsrude helped with their valuable suggestions in improving the methodology.
A study described in Chapter 8 has been published in:


The contribution of the author was in developing, implementing, and evaluating the method. Dr. Rasoulian helped with the developing of the statistical model. R. Bosma, R. Tong, Dr. Hollenstein and Dr. Harkness collected the MRI dataset. Dr. Abolmaesumi and Dr. Johnsrude helped with their valuable suggestions in improving the methodology.
Table of Contents

Abstract ................................................................. ii
Preface ................................................................. iv
Table of Contents .................................................... vii
List of Tables .......................................................... x
List of Figures .......................................................... xi
Glossary ................................................................. xiv
Acknowledgements ...................................................... xvii
Dedication ............................................................... xviii

1 Introduction .......................................................... 1
  1.1 Brain Imaging Data Analysis ................................... 1
  1.2 Classification of Individuals ................................... 2
  1.3 Multi-task or Multi-modal Analysis ............................ 3
  1.4 Proposed Framework ........................................... 5
     1.4.1 Objective .................................................... 5
     1.4.2 Contributions .............................................. 6
  1.5 Structure of Thesis ............................................. 6

2 Datasets .............................................................. 12
  2.1 Multi-task fMRI data: Speech Comprehension Study .......... 12
     2.1.1 Listening Study .......................................... 12
     2.1.2 Participants ............................................. 13
     2.1.3 Data Acquisition ........................................ 13
     2.1.4 Data Preprocessing ..................................... 13
  2.2 Structural MRI data: Major Depressive Disorder ............. 14
     2.2.1 Participants ............................................. 14
# Table of Contents

2.2.2 Clinical Examination ................................. 14  
2.2.3 Behavioural Results ................................. 15  
2.2.4 Data Acquisition ................................. 16  

3 Fusion Analysis of Functional MRI Data for Classification of Individuals ................................. 17  
3.1 Introduction ........................................... 17  
3.2 Method ................................................. 19  
   3.2.1 Joint Independent Component Analysis (jICA) ........................................... 20  
   3.2.2 Selection of Optimal Joint Sources ........................................... 21  
   3.2.3 Automatic Classification of Young and Older Subjects ........................................... 22  
3.3 Results ................................................. 23  
   3.3.1 Statistical Difference Among Joint Sources ........................................... 24  
   3.3.2 Selection of Optimal Joint Sources ........................................... 24  
   3.3.3 Classification ........................................... 24  
3.4 Discussion ................................................. 29  

4 Joint Sparse Representation Analysis ................................. 35  
4.1 Introduction ........................................... 35  
4.2 Method ................................................. 36  
4.3 Simulations .............................................. 40  
   4.3.1 Data Generation ........................................... 40  
   4.3.2 Experiments and Results ........................................... 43  
4.4 Results ................................................. 46  
4.5 Discussions .............................................. 48  

5 Reliability Analysis and Visualization ................................. 52  
5.1 Introduction ........................................... 52  
5.2 Method ................................................. 55  
   5.2.1 Preprocessing ........................................... 56  
   5.2.2 Sparse Representation Analysis ........................................... 56  
   5.2.3 Reliability Analysis ........................................... 56  
   5.2.4 Visualization using t-SNE ........................................... 57  
5.3 Simulations .............................................. 58  
   5.3.1 Data Generation ........................................... 58  
   5.3.2 Experiments and Results ........................................... 59  
5.4 Experiments and Results ........................................... 63  
5.5 Discussions .............................................. 64
# Table of Contents

6 Multi-object Statistical Analysis of Major Depressive Disorder .......................... 67
   6.1 Introduction ........................................... 67
   6.2 Method .............................................. 70
      6.2.1 Preprocessing ................................. 70
      6.2.2 Pose and Shape Analysis ...................... 72
      6.2.3 Statistical Analysis ......................... 73
   6.3 Results ............................................. 75
      6.3.1 Volume Analysis .............................. 75
      6.3.2 Pose and Shape Analysis ...................... 77
   6.4 Discussion ........................................ 81

7 Fusion Analysis of Brain Shape Deformations and Local Composition of Tissue ........ 84
   7.1 Introduction ........................................ 84
   7.2 Method .............................................. 87
      7.2.1 Features ..................................... 87
      7.2.2 Joint Independent Component Analysis .......... 88
   7.3 Experiments and Results ............................ 90
   7.4 Discussion ........................................ 95

8 Simultaneous Analysis of Pose, Shape, and Tissue Composition .......................... 99
   8.1 Introduction ........................................ 99
   8.2 Method .............................................. 101
      8.2.1 Feature Generation ......................... 101
      8.2.2 Group Joint Analysis ....................... 103
      8.2.3 Classification ............................... 103
   8.3 Results and Discussion ............................ 104

9 Conclusion and Future Work .................................................. 107
   9.1 Conclusion ........................................ 107
   9.2 Future Work ...................................... 108

Bibliography ......................................................... 111
List of Tables

3.1 TalaRaich coordinates for the most discriminative source map in three contrasts. .................................................. 27
4.1 ROIs used for each group for each task. .............................. 42
4.2 Similarity coefficients showing the result of jICA and jSRA methods for different combinations of tasks. ......................... 45
5.1 ROIs used for each group to create the simulated data. ......... 59
5.2 Normalized AUC for the four estimated components using the sparse analysis and ICA methods. .............................. 64
6.1 Normalized pose parameters of brain structures. L and R show the assigned anatomical left and right hemispheres. .... 77
7.1 MNI coordinates for the most discriminative source map in three contrasts. .................................................. 94
7.2 P-values of the most significant joint source. .................... 95
7.3 Classification error obtained from discriminant analysis of the mixing coefficients generated by jICA. ......................... 95
List of Figures

1.1 Proposed joint source-based analysis (jSBA) framework for classification of individuals. ......................... 11
2.1 Distribution of BDI for the control and depressed groups. . . 15
3.1 Schematic of the joint Independent Component Analysis (jICA) method. ............................................. 20
3.2 jICA of brain patterns. ................................................. 26
3.3 The effect of combining different contrasts on differentiation between histograms. ................................. 29
3.4 Classification accuracy for different number of features used. . 30
3.5 ROC curves and detection reliability resulting from four different numbers of features used. ..................... 31
3.6 ROC curves and detection reliability for different contrasts. . 32
3.7 ROC curves and detection reliability for different numbers of ICs. .......................................................... 33
4.1 Flowchart of the analysis, including preprocessing, feature selection, joint sparse representation analysis, and visualizing the output sources. ..................................................... 38
4.2 Schematic of the joint Source Representation Analysis (jSRA) method. .................................................... 39
4.3 ROIs representing the activated regions within the brain. . . 42
4.4 fMRI signal of a voxel before (a) and after randomization (b), the created synthetic activation signal (c), and simulated fMRI signal (d). ................................................................. 43
4.5 Average Jaccard (a), precision (b), and sensitivity (c) for combinations of two, three and four tasks. ............ 46
4.6 Significant joint source for jSRA (a-d) and jICA (f-h) methods, using 4 sources. ..................................... 47
4.7 Significant joint source for jSRA (a-d) and jICA (f-h) methods, using 8 sources. ..................................... 48
List of Figures

4.8 Significant joint source for jSRA (a-d) and jICA (f-h) methods, using 12 sources. 49

5.1 Flowchart of the analysis, including preprocessing, sparse representation analysis, reliability analysis, and visualization. 55

5.2 Reliability and visualization analysis of the simulated data for two different sets of parameters (first and second row). 61

5.3 Trustworthiness index as a function of neighborhood size. 62

5.4 Distribution of cophenet correlation coefficient in multiple runs of the visualization and clustering. 62

5.5 Activation maps obtained for each group using different parameter settings. 63

5.6 Reliability and visualization analysis of the simulated data for two different sets of parameters (first and second row). 65

6.1 Schematic of the pose and shape statistical analysis of multiple brain structures. 71

6.2 Segmented structures in both hemisphere of the brain which are used for multi-object statistical analysis. 72

6.3 Distribution of the volume of each structure between the two groups. 76

6.4 Shape principal component that was significantly different between the two groups. 78

6.5 Shape principal component that was significantly different between the two groups. 79

6.6 Pose (a) and shape (b) scores that generated the significant difference between the MDD subjects and controls across the beck Depression Inventory Index (BDI). 80

7.1 Joint Independent Component Analysis (jICA) of brain tissue composition and shape deformation. 93

7.2 Group-average histogram for the whole dataset on GM (a), WM (b), and deformation field (c). 96

7.3 Renyi divergence criteria values for different combination of features on differentiation between histograms. 97

8.1 Proposed joint source-based analysis (jSBA) framework for classification of individuals. 101

8.2 Schematic of the joint sparse representation method [150]. 102

8.3 Reliability and visualization analysis of the 45 subjects with and without depression. 105
8.4 Sparse coefficients for the significant joint sources for MDD subjects and healthy controls (a, b), respectively. . . . . . . . [106]
Glossary

ACC anterior cingulate cortex. 66
AUC Area Under Curve. 23
BDI Beck Depression Inventory. 14
CSF Cerebrospinal Fluid. 69
DARTEL Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra. 9
DBM Deformation Based Morphometry. 68
DF Deformation Fields. 85
DSM-IV-TR Diagnostic and Statistical Manual of Mental Disorders. 14
DTI Diffusion Tensor Imaging. 106
EEG Electroencephalography. 3
fMRI Functional Magnetic Resonance Imaging. 1
FN False Negative. 22
FP False Positive. 22
GLM General Linear Model. 3
GM Grey Matter. iii, 2, 69
Hc hippocampus. 66
ICA Independent Component Analysis. 3
<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITG</td>
<td>inferior temporal gyrus. xv, 71</td>
</tr>
<tr>
<td>jICA</td>
<td>joint Independent Component Analysis. ii</td>
</tr>
<tr>
<td>jSBA</td>
<td>joint Source Based Analysis. ii</td>
</tr>
<tr>
<td>jSRA</td>
<td>joint Sparse Representation Analysis. iii</td>
</tr>
<tr>
<td>K-SVD</td>
<td>K-Singular Value Decomposition. 34</td>
</tr>
<tr>
<td>LDA</td>
<td>Linear Discriminant Analysis. 2</td>
</tr>
<tr>
<td>loni</td>
<td>Laboratory of Neuro Imaging. 8</td>
</tr>
<tr>
<td>LPBA40</td>
<td>LONI Probabilistic Brain Atlas. 69</td>
</tr>
<tr>
<td>MDD</td>
<td>Major Depressive Disorder. 6</td>
</tr>
<tr>
<td>MDL</td>
<td>Minimum Description Length. 21</td>
</tr>
<tr>
<td>MEG</td>
<td>Magnetoencephalography. 3</td>
</tr>
<tr>
<td>MNI</td>
<td>Montreal Neurological Institute. 8</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnetic Resonance Imaging. 3</td>
</tr>
<tr>
<td>MTG</td>
<td>middle temporal gyrus. xv, 71</td>
</tr>
<tr>
<td>OMP</td>
<td>Orthogonal Matching Pursuit. 39</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal Component Analysis. 3</td>
</tr>
<tr>
<td>PVD</td>
<td>Provoked Vestibulodynia. 107</td>
</tr>
<tr>
<td>RFA</td>
<td>Recursive Feature Addition. 3</td>
</tr>
<tr>
<td>ROC</td>
<td>Receiver Operating Characteristic. 23</td>
</tr>
<tr>
<td>sMRI</td>
<td>Structural Magnetic Resonance Imaging. iii, 1</td>
</tr>
<tr>
<td>STG</td>
<td>superior temporal gyrus. xv, 66, 71</td>
</tr>
<tr>
<td>SVD</td>
<td>Singular Value Decomposition. 3</td>
</tr>
</tbody>
</table>
SVM Support Vector Machine. 2
TBM Tensor Based Morphometry. 68
TN True Negative. 22
TP True Positive. 22
TR Repetition Time. 13
VBM Voxel Based Morphometry. 1
VVS Vulvar Vestibulitis Syndrome. 107
WM White Matter. iii, 69
Acknowledgements

I offer my enduring gratitude to the faculty, staff and my fellow students at UBC and Queen’s University, who have inspired me to continue my work in this field. I owe particular thanks to my supervisors, Dr. Purang Abolmaesumi and Dr. Ingrid Johnsrude for their guidance, their scientific as well as personal support, and dedication to research.

A number of faculty members and staff in Department of Electrical and Computer Engineering, UBC, Department of Radiology, and Department of Psychology, Queen’s University have had a significant role in my research. I would like to thank and appreciate contributions from Dr. Kate Harkness, Dr. Tom Hollenstein, Dr. Caroline Pukall, Dr. Roger Tam, Kris Marblek, Dr. Amir Tahmasebi, Dr. Kate Sutton, Heather MacDonald, and Rachael Bosma. Special thanks to Dr. Abtin Rasoulian and Saman Nouranian for their insightful feedback and sharing their knowledge in medical image analysis, and software development.

I would like to thank the Natural Sciences and Engineering Research Council of Canada (NSERC), the Canadian Institutes of Health Research (CIHR), and UBC for funding this work.

Special thanks are owed to my parents, who have supported me throughout my years of education, both morally and financially. Last but not least I would like to thank my wife, Marjan, who has been with me in every step with her support, encouragement, quiet patience and unwavering love.
Dedication

I would like to dedicate my thesis to my beloved parents, Maryam and Mahmoud, and my wife, Marjan.
Chapter 1

Introduction

1.1 Brain Imaging Data Analysis

Many neurological disorders are associated with changes in brain structures or patterns of activity that can be observed using functional and anatomical neuroimaging. Over the past two decades, functional and structural Magnetic Resonance Imaging (fMRI and sMRI) have been used to identify regions of activity, to determine volume, shape, and position of brain structures, to diagnose diseases and lesions, and also for neurological and cognitive psychology research.

fMRI studies are typically analyzed to reveal regional specialization for cognitive functions or tasks, or to compare patterns of activity between two groups, such as patients and healthy control participants [43]. Usually simple comparisons of conditions are performed, and contrasts of interest are created. These contrasts reveal the regions showing the most consistent effects, and regions that differ consistently between groups [73, 75, 76]. Marked variability within groups can make it difficult to determine whether groups differ reliably with respect to the localization and extent of the activation. Furthermore, these approaches, in which all comparisons are considered independently, are not sensitive to shared information among different tasks on which functional imaging data are collected.

sMRI studies have revealed the neuroanatomical correlates of neurological disorders, characterizing differences in shape, volume, and local tissue composition. Proposed approaches are classified into those that measure differences in brain shape, and those that measure differences in the local volume (and concentration) of brain tissue after macroscopic differences in shape have been discounted [6]. The former approaches use the deformation fields that map any individual brain onto some standard reference to characterize neuroanatomy. The latter approaches, such as Voxel Based Morphometry (VBM) [6], focus on the local composition of brain tissue, and compare images for their tissue composition on a voxel-by-voxel basis after the deformation fields have been used to spatially normalize the images [13]. In other words, conventional computational neuroanatomic techniques have
either used the deformation fields themselves to characterize brain structural variation, or have used these fields to normalize images that are then entered into an analysis of regionally specific differences. Ideally, a procedure like VBM should be able to automatically identify any structural abnormalities in a single Grey Matter (GM) concentration image. However, even with many hundreds of subjects in a database, the method may not be powerful enough to detect subtle abnormalities in individuals [6]. A more powerful procedure might be to use some form of voxel-wise multivariate approach. Within a multivariate framework, in addition to images of GM concentration, other image features such as White Matter (WM) concentration, and information from the spatial normalization procedure could be included [6].

1.2 Classification of Individuals

Our understanding of the brain basis of disorders has benefited greatly during the past decade from important advances in machine learning and classification techniques. Given the predicted increase in prevalence of brain disorders including developmental, psychiatric and neurodegenerative diseases in the coming decades [160, 61], early detection and intervention in persons with mild brain impairment is of great importance. To better characterize patients with early impairment, it is critical to develop tools that can be used in reliable classification of patients in their early stage of the disease. These classification tools will initially be used in clinical research and ultimately be use in the clinical treatment of patients at risk for brain disorders [160]. Generally, classification based on neuroimaging data is not trivial, due to the high dimensionality of input feature space and small set of subjects that is usually available.

Although scant previous work deals with the specific problem of classification of individuals based on brain imaging data, the closely related goal, of using machine learning to decode stimuli, mental states, and behaviors from fMRI data is rapidly gaining in popularity; particularly the set of methods called representational similarity analysis, or pattern-information analysis (see [132] [145] for tutorial reviews). In this context, Haxby and colleagues showed that fMRI activation patterns are different when viewing a photograph of a face from viewing a house, a shoe, or a chair [88]. Using a similar dataset, Cox and Savoy successfully classified patterns of fMRI activation evoked by the presentation of photographs of various categories of objects, by applying Support Vector Machine (SVM) and Linear Discriminant Analysis (LDA) [48]. Mitchell et al. successfully trained classifiers to
1.3 Multi-task or Multi-modal Analysis

automatically decode the subject’s cognitive state at a single time instant or interval [129]. De Martino et al. combined multivariate voxel selection and SVM for classification of fMRI spatial patterns [121]. Pereira combined dimensionality reduction and classification into a single learning objective to achieve better learning performance [144]. Unlike these studies that used relatively simple stimuli or images drawn from fixed categories, Kay et al. used natural receptive-field models to identify a specific image, viewed by an observer, from a large set of natural images [95]. Their group further combined structural and semantic encoding models, and prior information about the structure and semantic content of natural images, to produce accurate reconstructions of observed natural images from brain activities [133]. Schrouff et al. used feature extraction methods with different classifiers to decode semi-constrained brain activity patterns, where number and duration of mental events were not externally imposed [171].

In this thesis, the goal is to characterize brain structural and functional changes and to use it for classification of individuals, rather than to detect transient cognitive states. This procedure typically consists of three steps. In functional MRI analysis, the first step is to determine the activation maps using a data-driven method such as group-Independent Component Analysis (group-ICA) [27], or a model-based approach such as General Linear Model (GLM) [69]. In structural MRI analysis, the first step is to determine brain tissue composition maps [13], or brain deformation maps [6]. The second step is to reduce the dimensionality of the data and compute representative features using Principal Component Analysis (PCA) [58, 70], Singular Value Decomposition (SVD) [3], Independent Component Analysis (ICA) [69], GLM or Recursive Feature Addition (RFA) [171]. In the third step, a classification is performed on the obtained features. These approaches have only focused on the classification based on single comparison of conditions or structural maps, and are not sensitive to shared information among different contrasts generated from those comparisons.

1.3 Multi-task or Multi-modal Analysis

Recently, considerable attention has been focused on combining data across multiple modalities or tasks, to provide knowledge of the joint information that may exist among those sources [182]. These analyses may be important to better understand complex disorders that affect many aspects of the brain (such as its structure, function, and organization; [30]). The premise of multi-modality approaches is that each imaging modality provides comple-
1.3. Multi-task or Multi-modal Analysis

ementary information about different tissue characteristics, and at different spatial and temporal resolutions. Many techniques have been proposed to combine multi-modal or multi-task fMRI information. These techniques can be categorized into two main types: data-integration and data-fusion methods [29, 52, 168]. Data-integration techniques use one imaging modality to improve the results of another modality (for example, registration of EEG or MEG to MRI [85]; and the use of fMRI to estimate the location of dipoles or the distribution of neural sources prior to EEG [109]). On the other hand, data fusion techniques utilize multiple modalities [30] or tasks [31] to take advantage of combined information. Generally, due to weak cross-modality relationships and intersubject variability, finding one-to-one correspondence in multi-modal images is difficult; however, performing fusion analysis across multiple subjects makes this an easier problem to solve. In this type of analysis, each modality is usually reduced to a feature that is a lower-dimensional representation of a selected brain structure or task-related activation pattern. Using variations across individual subjects, associations across the features can be explored [29]. Joint Independent Component Analysis (jICA) [30] is one multivariate technique for fusion analysis. jICA is a group-level analysis technique that uses extracted features from individual subjects’ data (i.e. multiple modalities or functional contrast images) and tries to maximize the independence among joint components.

Classification based on multi-modal and/or multi-task information may improve our knowledge of neurological disorders that affect multiple aspects of brain structure, function and networks. However, to the best of our knowledge, prior studies did not show the effect of combining datasets across tasks or modalities to increase the sensitivity of the classification for neuroimaging applications, where usually small number of subjects are available. jICA represents a powerful way to reduce the dimensionality of fMRI data sets, permitting classification of subjects based on different functional and structural features. jICA and its extensions have been successfully applied to study aphasia [180] and schizophrenia [30, 31, 108, 182, 183]; however, jICA data have not been used before to quantitatively classify individuals as belonging to one group or another.

Although results of previous studies have shown that ICA-based joint analysis techniques, such as joint ICA [30], parallel ICA [108], and coefficient-constrained ICA [183], accurately identify sources of common variance among features, the theoretical assumption of independence of the functional patterns extracted by ICA algorithms is not guaranteed in practice, and components are separated on the basis of spatial sparsity rather than dependence [51]. Therefore, other mathematical properties of brain fMRI data
1.4. Proposed Framework

Diagnosis and clinical management of neurological disorders that affect brain structure, function and networks would benefit substantially from the development of new techniques that combine multi-modal and/or multi-task information. Here, we propose a joint Source-based Analysis (jSBA) framework to identify common information across different functional and structural features in data from fMRI and sMRI experiments, for classification of individuals with neurological and psychiatric disorders. The framework consists of three components, as shown in Fig. 1.1: 1) individual’s feature extraction, 2) joint group analysis, and 3) classification of individuals based on the group’s extracted features. In each component, combinations of novel and state of the art methods have been used. In the proposed framework, information from multi-modal and/or multi-task datasets is reduced to a feature that is a lower-dimensional representation of a selected brain structure or task-related activation pattern. For each individual, features are used within a source-based analysis method to generate basis brain sources and their corresponding modulation profiles, used to classify individuals into different categories.

1.4.1 Objective

The global objective of this thesis is to propose a framework that can distinguish patients and healthy controls where the number of available subjects are low, and the between group differences are subtle. To this end, we propose to use multi-modal and/or multi-task brain imaging data, and take advantage of the complementary information that exist among the modalities or tasks for group classification. We investigate the joint analysis of multiple features, some of which previously have not been integrated or fused. We further develop techniques that are appropriate for joint analysis of brain imaging datasets. As a corollary to this objective, we postulate that the proposed framework can identify joint information across different structural and functional features.
1.4.2 Contributions

This study develops a framework for classification of individuals based on fusing brain imaging information. In the course of achieving this objective, the following contributions were made:

- Investigating the use of multi-task fMRI for classification of individuals.
- Proposing a novel joint sparse representation technique for joint group analysis of multi-task fMRI data.
- Proposing a new technique for reliability analysis and visualization of sparse representation algorithms.
- Investigating the use of morphometric analysis of multiple brain structures for classification of adolescents with and without depression.
- Proposing a novel way to combine brain structural information such as shape, pose, and tissue composition within the framework.

1.5 Structure of Thesis

The rest of this thesis is subdivided into seven chapters as outlined below:

Chapter 2: Materials

In this chapter we describe the two datasets (functional and structural MRI data) which have been used to evaluate the proposed framework. The experimental fMRI data was acquired from sixteen young (age: 19-26) and sixteen older (age: 57-73) adults obtained from multiple speech comprehension tasks within subjects. We will use this dataset in chapters 3, 4, and 5. The structural MRI data was acquired from 16 females (aged 16 to 21) and 3 males (aged 18) with early-onset Major Depressive Disorder (MDD), and 25 female and 1 male healthy control participants, drawn from the same age range. This dataset will be used in chapters 6, 7, and 8.

Chapter 3: Fusion Analysis of Functional MRI Data for Classification of Individuals

Classification of individuals based on patterns of brain activity observed in functional MRI contrasts may be helpful for diagnosis of neurological disorders. Prior work for classification based on these patterns have primarily
focused on using a single contrast, which does not take advantage of complementary information that may be available in multiple contrasts. Where multiple contrasts are used, the objective has been only to identify the joint, distinct brain activity patterns that differ between groups of subjects; not to use the information to classify individuals. In this chapter, we use joint Independent Component Analysis (jICA) within a Support Vector Machine (SVM) classification method, and take advantage of the relative contribution of activation patterns generated from multiple fMRI contrasts to improve classification accuracy. Young (age: 19-26) and older (age: 57-73) adults (16 each) were scanned while listening to noise alone and to speech degraded with noise, half of which contained meaningful context that could be used to enhance intelligibility. Functional contrasts based on these conditions (and a silent baseline condition) were used within jICA to generate spatially independent joint activation sources and their corresponding modulation profiles. Modulation profiles were used within a non-linear SVM framework to classify individuals as young or older. Results demonstrate that a combination of activation maps across the multiple contrasts yielded an area under ROC curve of 0.86, superior to classification resulting from individual contrasts. Moreover, class separability, measured by a divergence criterion, was substantially higher when using the combination of activation maps.

Chapter 4: Joint Sparse Representation Analysis

Prior research using multi-task analysis in fMRI, such as the proposed approach in Chapter 3, has mainly assumed that brain activity patterns evoked by different tasks are independent. This may not be valid in practice. In this chapter, we use sparsity, which is a natural characteristic of fMRI data in the spatial domain, and propose a joint Sparse Representation Analysis (jSRA) method to identify common information across different functional contrasts in data from a multi-task fMRI experiment. Sparse representation methods do not require independence, or that the brain activity patterns be non-overlapping. We use functional contrast images within the joint sparse representation analysis to generate joint activation sources and their corresponding sparse modulation profiles. We evaluate the use of sparse representation analysis to capture individual differences with simulated fMRI data and with experimental fMRI data. The same experimental fMRI data as in Chapter 3 was used, where an independent measure (namely, age in years) can be used to differentiate between groups. Simulation results show that this method yields greater sensitivity, precision and higher Jaccard indices (which measures similarity and diversity of the true and estimated brain ac-
1.5. Structure of Thesis

tivation sources) than does the jICA method. Moreover, superiority of the jSRA method in capturing individual differences was successfully demonstrated using experimental fMRI data.

Chapter 5: Reliability Analysis and Visualization

Sparse representation analysis of neuroimaging data has been shown to be effective for detection of functional activation, for identification of brain functional networks, for multivariate pattern analysis and as shown in Chapter 4, for classification of individuals. However, results of a sparse analysis should be interpreted cautiously, because they may vary over multiple runs of the algorithm and depend on the initialization, parameter values and optimization algorithms employed. In this chapter, we propose a way to assess the reliability of such analyses, and to visualize the brain patterns obtained from a sparse representation analysis. We run the sparse analysis multiple times for each parameter value, and cluster the estimated components. The clusters are nonlinearly mapped into a low-dimension space, which enables further interpretation of the components, and identification of the best parameters values. We evaluate the use of the proposed approach using both simulated and experimental fMRI data. Results show that we can successfully identify reliable components and select the best parameters using the proposed approach.

Chapter 6: Multi-object Statistical Analysis of Major Depressive Disorder

In the next three chapters we use the structural MRI dataset of individuals with and without depression. Major depressive disorder (MDD) has previously been linked to structural changes in several brain regions, particularly in the medial temporal lobes [14, 15]. This has been determined using voxel-based morphometry, segmentation algorithms, and analysis of shape deformations [13, 16, 147, 196, 204]; these are methods in which information related to the shape and the pose (the size, and anatomical position and orientation) of structures is lost. In this chapter, we incorporate information about shape and pose to measure structural deformation in adolescents and young adults with and without depression (as measured using the Beck Depression Inventory and Diagnostic and Statistical Manual of Mental Disorders criteria). We focus on changes in cortical and subcortical structures, and use a multi-object statistical pose and shape model to analyze imaging data from 16 females (aged 16 to 21) and 3 males (aged 18) in with early-
onset MDD, and 25 female and 1 male healthy control participants, drawn from the same age range. Hippocampus, parahippocampal gyrus, putamen, and the superior, inferior and middle temporal gyri in both hemispheres of the brain are automatically segmented using the loni Probabilistic Brain Atlas [176] in MNI space. Points on the surface of each structure in the atlas are extracted and warped to each participant's structural MRI. These surface points are analyzed to extract the pose and shape features. Pose differences are detected between the two groups, particularly in the left and right putamen, right hippocampus, and the left and right inferior temporal gyri. Shape differences are detected between the two groups, particularly in the left hippocampus and in the left and right parahippocampal gyri. Furthermore, pose measures are significantly correlated with BDI score across the whole (clinical and control) sample. Since the clinical participants were experiencing their very first episodes of MDD, morphological alteration in the medial temporal lobe appears to be an early sign of MDD, and is unlikely to result from treatment with antidepressants. Pose and shape measures of morphology, which are not usually analyzed in neuromorphometric studies, appear to be sensitive to depressive symptomatology.

Chapter 7: Fusion Analysis of Brain Shape Deformations and Local Composition of Tissue

Computational neuroanatomical techniques, such as the one proposed in Chapter 6, that are used to evaluate the structural correlates of disorders in the brain typically measure regional differences in grey matter or white matter, or measure regional differences in the deformation fields required to warp individual datasets to a standard space. Our aim in this chapter is to combine measurements of regional tissue composition and deformations in order to characterize a particular brain disorder (here, major depressive disorder). We use structural magnetic resonance imaging (MRI) data from young adults in a first episode of depression, and from an age- and sex-matched group of non-depressed individuals. After DARTEL groupwise registration, we obtained grey matter (GM) and white matter (WM) tissue maps in the template space, along with the deformation fields required to warp the DARTEL template to the GM and WM maps in the population. These three features, reflecting tissue composition and shape of the brain, are used within jICA to extract spatially independent joint sources and their corresponding modulation profiles. Coefficients of the modulation profiles are used to capture differences between depressed and non-depressed groups. The combination of hippocampal shape deformations and local composition
of tissue (but neither shape nor local composition of tissue alone) is shown to discriminate reliably between individuals in a first episode of depression and healthy controls, suggesting that brain structural differences between depressed and non-depressed individuals do not simply reflect chronicity of the disorder but are there from the very outset.

Chapter 8: Simultaneous analysis of Pose, Shape, and Tissue Composition

In this chapter we use the jSBA framework to combine the two previous chapters, which were shown to be effective in identification of brain structural variations in patients with Major Depressive Disorder (MDD). In this framework, features representing position, orientation and size (i.e. pose), shape, and local tissue composition are extracted. Subsequently, simultaneous analysis of these features within a joint analysis method is performed to generate the basis sources that show significant differences between subjects with MDD and those in healthy control. Moreover, in a cross-validation leave-one-out experiment, we use a random forest classifier to identify individuals within the MDD group. Results show that we can classify the MDD subjects with an accuracy of 67% solely based on the information gathered from the joint analysis of pose, shape, and tissue composition in multiple brain structures.

Chapter 9: Conclusion and Future Work

This chapter includes a short summary followed by suggestions for future work.
1.5. Structure of Thesis

Figure 1.1: Proposed joint source-based analysis (jSBA) framework for classification of individuals.
Chapter 2

Datasets

The proposed framework has been used for analysis of brain activity patterns in a speech comprehension task, and for analysis of depression in adolescence and young adulthood, to demonstrate its potential to classify groups of subjects based on functional and structural datasets. In the next two sub-sections we briefly describe the datasets used for the two applications.

2.1 Multi-task fMRI data: Speech Comprehension Study

2.1.1 Listening Study

Subjects were asked to listen to sentences in the scanner and try to understand them as well as they could. Sentences with and without coherent sentence-level meaning ('coherent' and 'anomalous' sentences, respectively) were taken from those used by [55] and were mixed with noise that had the same long term spectrum of the speech and the amplitude envelope of the signal to be masked (Signal-Correlated Noise: SCN; [170]) at six different signal to noise ratios (SNRs): -5 dB, -3.5 dB, -2.5 dB, -1 dB, 0 dB, and 2.5 dB. Clear speech was also tested, making 7 sentence conditions. Coherent and anomalous sentences were divided into 7 sets, which were pseudorandomly assigned to conditions such that each sentence set was tested in each of the seven SNR conditions (including clear speech) an equal number of times, across participants. Over the scanning session, each participant heard 14 trials of each sentence type at each SNR, half of which were followed by 'repeat' trials requiring the participant to repeat as much of the sentence as possible. The stimuli assigned to repeat trials were counterbalanced across participants, and repeat and non-repeat trials were randomly intermixed for each participant. Intelligibility, defined here as the proportion of words correctly reported, was obtained for each signal quality level and for each sentence type, for each participant, from the repeat trials. Moreover, data from 14 trials of SCN on its own and 16 trials of silence were obtained. There were 324, trials distributed across four 81-trial sessions of testing.
2.1. Multi-task fMRI data: Speech Comprehension Study

2.1.2 Participants
Sixteen young (mean age: 21.1, range: 19-26, 11 female) and sixteen older (mean age: 64.2, range: 57-73, 11 female) adults were scanned. All subjects were native speakers of English, without any history of neurological illness, head injury, or hearing impairment. This study was cleared by the Queen’s University Health Sciences Research Ethics Board, and written informed consent was obtained from all participants.

2.1.3 Data Acquisition
The fMRI data were acquired using a 3.0 Tesla Siemens Trio MRI system with a 12-channel head coil in the MRI facility at Queen’s University, Kingston, Canada. Each acquisition consisted of 32 contiguous slices with 4 mm thickness, field of view $211 \times 211$ mm, in plane resolution of $3.3 \times 3.3$ mm, resulting in a grid of $64 \times 64 \times 32$ voxels, each $3.3 \times 3.3 \times 4$ mm in volume. The Repetition Time (TR) was 9 sec and the acquisition time was 2 sec. This sparse GE-EPI imaging technique allowed for stimuli to be presented in the silent gaps between scans. Total functional imaging time was 48 minutes. Auditory stimuli and the visual ‘repeat’ instructions were presented to the participants using E-Prime v.1.2 and a NEC LT265 DLP projector. Participants viewed the screen via a mirror system mounted on the head coil [115].

2.1.4 Data Preprocessing
Before preprocessing, the Siemens motion correction algorithm (Available online at: http://imaging.mrc-cbu.cam.ac.uk/imaging/DataDiagnostics) was applied to the DICOM MR images, and then the DICOM images were converted to NIFTI format. The fMRI data were preprocessed using Statistical Parametric Mapping software (SPM8, Wellcome Department of Cognitive Neurology, London, UK). Preprocessing steps included realignment [74], coregistration [44] and the segmentation-based spatial normalization [7] of SPM8. The data were spatially smoothed using an 8-mm Gaussian kernel. Spatial smoothing has been previously shown to be effective at increasing functional signal-to-noise in SPM-based analyses [187]. The first scan of each session was discarded, and the rest were coded according to the auditory condition of the preceding stimulus and entered into a single-subject general linear model. The Finite Impulse Response (FIR) set was selected as the hemodynamic response function. Three functional contrasts were calculated: SCN versus silence, identifying brain regions that process the
2.2 Structural MRI data: Major Depressive Disorder

2.2.1 Participants

Nineteen depressed subjects (age: 18.1 ± 1.1, 3 male, all right-handed) who met DSM-IV-TR (Diagnostic and Statistical Manual of Mental Disorders; American Psychiatric Association, 2000) criteria for a current episode of MDD were recruited through referrals from community mental health clinics. Twenty-six healthy participants (age: 17.96 ± 0.2, 1 male, all right-handed) with no psychiatric history were recruited through community advertisement. Participants were excluded if they met current or lifetime criteria for bipolar disorder, a psychotic disorder, attention-deficit/hyperactivity disorder, a developmental disability (e.g., autism spectrum disorder), or a medical disorder that could cause depression (e.g., hypothyroidism). All participants were medication-free prior to the study. This study was cleared by the Queen’s University Health Sciences Research Ethics Board, and written informed consent was obtained from all participants and by a parent or guardian for participants under 18 years. All participants were compensated $10 for their time and were given a picture of their brain to keep.

2.2.2 Clinical Examination

All participants in the depressed group were diagnosed based on a structured diagnostic interview administered by an advanced doctoral student in clinical psychology (the Child and Adolescent version of the Schedule for Affective Disorders and Schizophrenia; K-SADS; [94]). The K-SADS is the gold standard for deriving DSM-IV-TR diagnoses in children and adolescents and is the most widely used measure for this purpose in clinical research. This clinician interview was administered by graduate-level students in clinical psychology who were trained and supervised by a registered clinical psychologist with over 20 years of expertise in the assessment and diagnosis of depression in adolescence. Participants were scored in the mild to severe depression range, as defined by a Beck Depression Inventory (BDI-II; [11]) score. The BDI is a 21-item measure designed to assess the presence and severity of depression symptoms, and is the most commonly used depression
2.2. Structural MRI data: Major Depressive Disorder

Figure 2.1: Distribution of BDI for the control and depressed groups. The central red mark is the median, the edges of the blue box are the 25th and 75th percentiles, and the whiskers show the extreme values of the volumes.

measure in adolescent and adult samples [50]. The BDI was administered by trained graduate or senior undergraduate-level students who went over the measure with each participant to ensure that they understood the questions. We chose not to include the Hamilton Depression Rating Scale and to focus exclusively on the BDI as an index of depression severity for the primary reason that there is evidence that the Hamilton possesses a poor psychometric profile [9].

2.2.3 Behavioural Results

The groups were matched for age ($p = 0.52$). There was no socioeconomic status (SES) differences between the subjects in the two groups ($p = 0.93$). The BDI differed significantly between the two groups ($p < .0001$). Fig. 2.1 shows the boxplot of the BDI for the two groups.
2.2. Structural MRI data: Major Depressive Disorder

2.2.4 Data Acquisition

The MRI data were acquired using a 3.0 Tesla Siemens Trio MRI scanner with a 12-channel head coil in the MRI facility at Queen’s University, Kingston, Canada. A whole-brain 3D MPRAGE T1-weighted anatomical image was acquired for each participant (voxel resolution of $1.0 \times 1.0 \times 1.0$ mm, flip angle $\alpha = 9^\circ$, $TR = 1760$ ms, and $TE = 2.6$ ms). The subjects filled out the BDI immediately after being in the scanner.
Chapter 3

Fusion Analysis of Functional MRI Data for Classification of Individuals

3.1 Introduction

Functional Magnetic Resonance Imaging (fMRI) studies are typically analyzed to reveal regional specialization for cognitive functions or tasks, or to compare patterns of activity between two groups, such as patients and normal-control participants [43]. Usually, simple comparisons of conditions are performed to reveal regions that are reliably active by the task of interest, and/or regions that differ reliably between groups [73, 75, 76]. Several studies have taken advantage of these identified regions for group classification based on fMRI data [3, 58, 70]. However, these approaches have only focused on the classification based on single comparison of conditions, and are not sensitive to shared information among different contrasts generated from those comparisons.

Many techniques have been proposed to combine multi-task fMRI information. Joint Independent Component Analysis (jICA) [30] is one multivariate technique for fusion analysis. It is an extension of ICA that combines information from multiple modalities or functional contrasts. The simplified noise-free ICA model seems to be sufficient for many applications [92], and has been successfully applied to fMRI data [28, 123, 183]. In this simplified model, the ICA components that contribute the least, and which may have a "speckled" spatial distribution in contrast images, are noise of unknown origin [123]. ICA is typically used as a first-level data-driven approach to find spatially or temporally independent brain sources of activity from fMRI

---

1This chapter is adapted from [148]: Mahdi Ramezani, Purang Abolmaesumi, Kristopher Marble, Heather Trang, and Ingrid Johnsrude. Fusion analysis of functional MRI data for classification of individuals based on patterns of activation, Brain Imaging and Behavior, 2013.
3.1. Introduction

scans of a person’s brain, while that person is performing a desired task [123]. Spatial ICA results in a set of independent components (spatial brain activation patterns), and a set of “mixing coefficients”.

Joint ICA is a group-level analysis technique that uses extracted features from individual subjects’ data and tries to maximize the independence among joint components. Assuming that the features, obtained from multiple modalities or multiple contrasts, share the same modulation profile (i.e. mixing-coefficient matrix), jICA uses more information to estimate the same number of mixing coefficients; and may therefore yield improved results compared to ICA. An additional advantage of jICA is the computation of modulation profiles along with the identified sources. Such profiles substantially reduce the dimensionality of the data, and can be used for group classification. The ability of jICA to reduce the number of dimensions is particularly important in the context of fMRI data analysis, where the dimensionality of input feature space is high and the number of available subjects is usually low. Although the discrimination ability of joint components has been investigated [184], to the best of our knowledge, the application of modulation profiles for classification of individuals has not, to our knowledge, been tried. Recently, Fan et al. used the modulation profile, which resulted from applying ICA to resting-state fMRI data, to classify individuals with schizophrenia and healthy controls [69]. Their work did not investigate the possibility of improved classification that could be obtained by combining multiple fMRI contrasts.

In this chapter of the thesis, we use jICA for group classification. We first identify the modulation profile (mixing coefficients) that reflects group differences by fusion analysis of multiple contrasts, and then use the resulting profile for group classification. We test this classification approach using fMRI data collected from 16 young and 16 older neurologically normal individuals who were scanned in multiple stimulus conditions in a speech perception experiment [55, 116]. The jICA modulation profile, which reflects group differences in activation patterns observed in three different functional contrasts, is used to automatically classify individuals as young or older. One contrast compares responses to unintelligible noise bursts, amplitude modulated with the temporal envelope of spoken sentences, with silence. A second contrast compares responses to sentences without coherent meaning (“anomalous” sentences, e.g., “Her good slope was done in carrot”) with the unintelligible noise bursts. A third contrast compares responses to sentences with coherent meaning (“coherent” sentences, e.g., “Her new skirt was made of denim”) to anomalous sentences. The intelligibility of anomalous sentences is determined by the quality of the signal, whereas the intelligi-
bility of coherent sentences is determined both by the quality of the signal, and by semantic knowledge. At any given signal quality, comprehension is greater for the latter than for the former (referred to hereafter as “context benefit”) [116]. Although older and younger adults do not differ in context benefit measured behaviorally [116], we examine whether patterns of activity in functional contrasts can be used to classify young and older people, without considering any information reflecting the distinct brain structural differences of the two groups [35] [81].

Joint ICA is used with data from the three contrasts to probe the unique and joint information among different contrasts and groups. First, joint independent components based on different combinations of these features, along with the mixing coefficients, are determined and then statistical differences among mixing coefficients (reflecting the network strengths) are examined using t-tests. Third, separability of the joint-source distributions is measured in order to assess the difference between distributions from different young and older participants [90]. Finally, the modulation profile extracted from the three functional contrasts is used to classify individuals as young or older, and the accuracy of this classification is assessed. We demonstrate that by fusing the three contrasts with jICA, the discrimination of subjects as young or older is substantially improved compared to using each individual contrast alone. Here we show the feasibility of this method by examining age-related differences in healthy subjects, where an independent measure (namely, age in years) can be used to differentiate between groups with 100% certainty. This "gold standard" allows us to validate the approach, which we expect to be applicable to real-world diagnostic problems without such a clear standard to differentiate groups.

3.2 Method

As mentioned in Section 2.1, the behavioural performance for each subject was measured as the words correctly reported at different SNRs. Pilot work revealed that in general at a given SNR older people reported fewer words than younger people. Accordingly we altered SNRs for the two groups to match behavioural performance. The average report score for anomalous sentences, which do not provide a contextual benefit, gives a good indication of low-level speech processing. A range of SNRs for each group was chosen in order to equate the behavioural performance while hearing the anomalous sentences. The ranges were -5 dB, -3.5 dB, -2.5 dB, -1 dB, and 0 dB for younger people and -3.5 dB, -2.5 dB, -1 dB, 0 dB, and +2.5 dB for older
3.2. Method

Joint Independent Component Analysis (jICA)

Joint ICA assumes a noise-free generative model \( X = AS \) where a source matrix \( S = [s_1, s_2, \ldots, s_N]^T \) combines with the mixing coefficients matrix \( A \) (also called the ICA loading parameters matrix) to generate the observations \( X = [x_1, x_2, \ldots, x_N]^T \). The \( j \)th row, \( s_j \), of \( S \) is the \( j \)th joint independent component, and \( M \) is the number of independent components. \( N \) is the number of participants and \( x_i \) is a vector formed by concatenating the brain features.

The jICA method, shown schematically in Fig. 3.1, involves finding \( U = WX \), where \( W = A^{-1} \) is called the unmixing matrix and \( U \) is the estimate of the joint source matrix \( S \). In this figure \( A_y, A_o \) indicate the submatrices associated with the young and old subjects. A MATLAB implementation of jICA is provided by the FIT 2.0b software [30], available online at http://mialab.mrn.org/software/.

Joint independent components are found using the Infomax algorithm [12], which is based on minimization of mutual information of components. In this
3.2. Method

The output entropy of a neural network is adaptively maximized with as many outputs as the number of Independent Components (ICs) to be estimated. The best way to estimate the most appropriate number of independent components is not clear. This number can affect the results of ICA, particularly if it is too small [112]. Information theoretic techniques, such as the Minimum Description Length (MDL) criterion [104], have been shown to be useful for selecting the number of brain basis patterns [104]. However, these techniques may not converge because of the heterogeneity in localization and size of individuals’ brain features.

The jICA procedure generates a set of joint independent components and associated mixing coefficients. These low-dimensional coefficients model the modulation of each subject’s feature by a joint source, and thus can be used as a criterion for capturing group differences. To investigate whether the groups were separable by different weightings of the joint sources, unpaired two-sample t-tests with unequal variance (heteroscedastic) on the mixing coefficients were performed. The z-scaled results indicated the joint components of interest.

3.2.2 Selection of Optimal Joint Sources

If two groups differ, then the distributions of their joint components should be separable [183, 29]. Separability can be quantified by computing a divergence between joint histograms. Group joint-sources are defined by $U_y = A_y^{-1}X_y = W_yX_y$, where $A_y$, $W_y$ and $X_y$ indicate the submatrices associated with the young subjects (similar for the older group). For each subject, the appropriate group joint source (e.g. $U_o$ for an older subject) was divided into three maps, which correspond to the three contrasts used in the jICA analysis. The map elements (each one representing a specific voxel) were sorted and thresholded, leaving a set of voxels statistically relevant to the joint source. Each voxel that survived thresholding in all three maps was included in a three-dimensional joint histogram in a bin defined by the three contrast values (from the input observation matrix $X$) at that voxel’s location.

The group-averaged joint histograms were then calculated by taking the mean of the joint histograms of each subject in the group. The difference between the two groups was then assessed using the Renyi divergence formula [90] with $\alpha = 0.5$ which has been shown to be optimal for discrimination.
3.2. Method

between pairs of close feature densities [90]:

\[
D_\alpha(P||Q) = \frac{1}{\alpha - 1} \log\left(\sum_{i=1}^{n} \frac{p_i^\alpha}{q_i^{\alpha - 1}}\right) = \frac{1}{\alpha - 1} \log\left(\sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha}\right) \tag{3.1}
\]

where \(P\) and \(Q\) are probability distributions, reflected in the group-averaged joint histograms.

The divergence is also computed for different combinations of contrasts. The higher the values of the Renyi divergence criterion, the better the discrimination between groups. Therefore, best combination of contrasts is the one that yields the highest divergence value.

3.2.3 Automatic Classification of Young and Older Subjects

In order to overcome the classification problems caused by high dimensionality of fMRI data and the small set of available subjects (16 in each group), columns of the mixing coefficients matrix, which reflect the weighting of each joint source in a subject’s contrast, were used as input features to a classification algorithm. A non-linear Support Vector Machine (SVM) was used to classify the subjects. SVM does not assume that data points conform to a specific model, but rather seeks to find the hyperplane that separates the two classes with maximum margin [189]. The hyperplane is defined by \(f(x) = \omega K_s(x) + b\), where \(K_s(x) = [k(x, s_1), ..., k(x, s_d)]\) is the vector of kernel functions centered at the support vectors, \(\omega\) is the parameter vector and \(b\) is a scalar. Radial Basis Functions (RBF) were used as the kernels: \(k(x, s) = \exp\left(-\frac{|x-z|^2}{\sigma^2}\right)\). The data were split into a training set and a test set. In the training phase, fusion analysis was repeated on only those subjects in the training set. This produced a new mixing coefficient matrix \(A_{(\text{train})}\) and joint source matrix \(S_{(\text{train})}\) that modeled the generation of the training observations \(X_{(\text{train})}\). The columns of the mixing matrix \(A_{(\text{train})}\) were used as input features to train the classifier, which divided the training subjects into two classes of young and older adults.

The input features for the test set, columns of \(A_{(\text{test})}\), were then found by least-squares solution of \(X_{(\text{test})} = A_{(\text{test})}S_{(\text{train})}\). The positions of these vectors in \(k\)-dimensional feature space, relative to the hyperplane found in running the classifier on the training set, determined the classification of each test subject. The number of columns of \(A\), or mixing coefficients, used in the classification was \(k\). A MATLAB implementation of the classifier provided by the Statistical Pattern Recognition Toolbox (STPRtool), available at http://cmp.felk.cvut.cz/cmp/cmp_software.html was used for this step.
Performance of the classification procedure was measured by repeatedly splitting the data into training and test sets and averaging classification performance across iterations. The splitting was done 200 times (each time selecting different 11 young and 11 older subjects as the training set and the remaining five subjects in each group as the test set). The False Positive (FP), False Negative (FN), True Positive (TP) and True Negative (TN) values were calculated and the ratio between TP and TN values to the total number of outcomes was taken as the performance metric. Selecting the significant features (mixing coefficients) is an important factor in classification results. The mixing coefficients were sorted by the p-values resulting from a two-sample t-test checking for differences between groups, and the k coefficients with the most significant difference were chosen for use in classification. Unlike some statistical classification methods, SVM does not provide posterior class probabilities ($P_i$). Without posterior probabilities, it is not possible to assess the performance of the classifier at other threshold values, and to measure the sensitivity and specificity of the classifier. Following the Platt [146] approach, we trained an SVM and later the parameters of an additional sigmoid function to map the values of SVM outputs to posterior probabilities. Using the posterior probabilities, the Receiver Operating Characteristic (ROC) curves were plotted and their Area Under Curve (AUC) calculated. The AUC metric is the most common way to compare the accuracy of classification methods in the machine learning community. Detection reliability $\rho$, was defined based on AUC as $\rho = 2 \times AUC - 1$ [152]. Under this definition, $\rho = 1$ for perfect detection and $\rho = 0$ for failure in detection.

The joint ICA classification result was compared to ICA for each of the three contrasts separately, to examine whether the fusion analysis has advantages over analysis of the results of each contrast separately. As in the fusion analysis, the mixing coefficients were employed as input features for classification of young and older adults.

### 3.3 Results

Although older adults need higher SNRs to achieve the same performance scores as young adults (i.e. they do not perform as well in noise), behaviourally there is no difference in the amount of benefit they get from contextual information compared to anomalous information. The goal of our fusion analysis is to examine whether jICA components can be used to accurately distinguish young and older adults on the basis of fMRI data from
3.3. Results

a speech perception experiment, despite the similarity in contextual benefit between the groups. The success of this analysis is evaluated by examining the statistical difference among the mixing coefficients of joint sources, by applying the Renyi divergence criterion, and by an automatic classification method. These tests are described in the following three subsections.

3.3.1 Statistical Difference Among Joint Sources

Unpaired two-sample t-tests (assuming unequal variance) were performed on the mixing coefficients, and two components were found to differ significantly (p-value = 0.00071 and p-value = 0.0301, number of subjects = 16 in each group) between the two groups. Fig. 3.2a, Fig. 3.2b, and Fig. 3.2c show the statistical Z-maps generated for the joint source (shown as rows of Map 1, Map 2, and Map 3 in Fig. 3.1) with the largest group difference. Fig. 3.2d shows that the mixing coefficients for this joint source have higher values in older subjects compared to younger subjects. Table 3.1 shows the corresponding stereotaxic coordinates in MNI space for this source.

3.3.2 Selection of Optimal Joint Sources

The sorted maximum Renyi divergence values for different combinations of contrasts are shown in Fig. 3.3. Higher values indicate better discrimination between the groups. It is clear that combining all three contrast images yields the best results. It can also be seen that the contrast comparing responses to anomalous sentences and to unintelligible noise is the most effective single contrast in separating the groups.

3.3.3 Classification

Fig. 3.4 shows the average classification accuracy for different numbers of features. Results show that the young and older subjects can be classified based on their patterns of activity across the three contrasts of interest. Considering the fact that the number of subjects is low and the dimensionality of the input fMRI dataset to the fusion framework is quite high, the results are very promising. By selecting only the first three features, an average classification accuracy of around 75% is obtained. Fig. 3.5 shows the ROC curves and detection reliability for four different numbers of features. Since each additional feature is (by definition) less important than the previous one, the addition of features beyond three leads to asymptotically improving discriminability.
3.3. Results

Using the first three features, the joint ICA classification result is compared to ICA for each of the three contrasts separately in Fig. 3.6, which shows the ROC curves, the detection reliability and the area under ROC curve obtained. Classification of individuals based on ICAs of the contrast comparing anomalous sentences and noise, and the contrast comparing SCN and silence, yield significantly (p-value < 0.05) better-than-chance classifications, but substantially lower than the result obtained by the jICA classification.
3.3. Results

Figure 3.2: Joint Independent Component Analysis (JICA) of brain patterns. The joint source map of the most significant component for the contrast of high vs anomalous sentences (a), for the contrast of anomalous sentences vs SCN (b), and for the contrast of SCN vs Silence (c), along with the mixing coefficients for the young and older subjects (d) is presented.
Table 3.1: Talaraich coordinates for the most discriminative source map in three contrasts. MNI coordinates of voxels, which are above a threshold of $|Z| > 3.5$, are converted to Talairach coordinates. L and R show the assigned anatomical left and right hemispheres, cc stands for cubic centimeters showing the volume concentration of voxels, the coordinates and value of the maximum $Z$ are also provided in the table. Not significant regions are shown by ns.

<table>
<thead>
<tr>
<th>Area</th>
<th>Brodmann Area</th>
<th>R/L volume (cc)</th>
<th>R/L random effects: Max Value (x, y, z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Higher-level patterns</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cingulate Gyrus</td>
<td>31</td>
<td>0.5/0.5</td>
<td>4.2($-4, -48, 35$)/4.3($7, -48, 35$)</td>
</tr>
<tr>
<td>Precuneus</td>
<td>7, 31, 39</td>
<td>0.4/0.5</td>
<td>4.2($-4, -48, 38$)/4.1($7, -48, 38$)</td>
</tr>
<tr>
<td>Posterior Cingulate</td>
<td>23, 31</td>
<td>0.2/0.2</td>
<td>3.9($0, -45, 31$)/3.8($3, -45, 31$)</td>
</tr>
<tr>
<td>Angular Gyrus</td>
<td>*</td>
<td>0.0/0.1</td>
<td>ns/3.6($41, -63, 40$)</td>
</tr>
<tr>
<td>Negative</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle Frontal Gyrus</td>
<td>*</td>
<td>0.0/0.2</td>
<td>ns/4.3($44, 15, 32$)</td>
</tr>
<tr>
<td>-</td>
<td>*</td>
<td>0.0/0.2</td>
<td>ns/4.0($48, 15, 32$)</td>
</tr>
<tr>
<td>Inferior Frontal Gyrus</td>
<td>9</td>
<td>0.0/0.3</td>
<td>ns/4.0($44, 15, 32$)</td>
</tr>
<tr>
<td>Precentral Gyrus</td>
<td>–</td>
<td>0.0/0.1</td>
<td>ns/3.7($48, 7, 39$)</td>
</tr>
<tr>
<td>Mid-level patterns</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle Temporal Gyrus</td>
<td>21, 22</td>
<td>0.7/1.6</td>
<td>4.8($-64, -5, -6$)/5.1($67, -29, 8$)</td>
</tr>
<tr>
<td>Middle Frontal Gyrus</td>
<td>6</td>
<td>0.1/1.0</td>
<td>3.6($-57, 10, 55$)/4.7($51, 10, 55$)</td>
</tr>
<tr>
<td>Superior Temporal Gyrus</td>
<td>21, 22, 42</td>
<td>0.3/1.7</td>
<td>3.8($-55, -32, 8$)/4.9($57, -32, 6$)</td>
</tr>
</tbody>
</table>

Continued on next page
### Table 3.1 – continued from previous page

<table>
<thead>
<tr>
<th>Area</th>
<th>Brodmann Area</th>
<th>R/L volume (cc)</th>
<th>R/L random effects: Max Value (x, y, z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>*</td>
<td>0.1/0.5</td>
<td>ns/4.2(63,−42,31)</td>
</tr>
<tr>
<td>Inferior Parietal Lobule</td>
<td>40</td>
<td>0.0/0.4</td>
<td>3.9(−51,−32,8)/4.6(57,−35,6)</td>
</tr>
<tr>
<td>Precentral Gyrus</td>
<td>6</td>
<td>0.0/0.1</td>
<td>ns/3.9(57,7,51)</td>
</tr>
<tr>
<td>Postcentral Gyrus</td>
<td>*</td>
<td>0.0/0.1</td>
<td>ns/3.6(63,−26,45)</td>
</tr>
<tr>
<td>Inferior Frontal Gyrus</td>
<td>*</td>
<td>0.0/0.1</td>
<td>ns/3.6(48,12,32)</td>
</tr>
<tr>
<td>Supramarginal Gyrus</td>
<td>40</td>
<td>0.0/0.1</td>
<td>ns/4.0(63,−43,28)</td>
</tr>
</tbody>
</table>

#### Lower-level patterns

**Positive**

- Postcentral Gyrus | 40, 43 | 1.1/0.1 | 6.0(−74,−16,20)/3.6(67,−24,23) |
- Superior Temporal Gyrus | 22, 41, 42 | 1.9/0.7 | 5.6(−74,−22,17)/4.4(67,−28,24) |
- Transverse Temporal Gyrus | 41, 42 | 1.2/0.2 | 5.5(−67,−10,12)/4.0(67,−13,13) |
- Precentral Gyrus | 43 | 0.1/0.0 | 4.7(−64,−7,12)/ns |
- Insula | 13 | 0.0/0.2 | ns/4.2(63,−31,24) |
- | * | 0.1/0.0 | 5.9(−78,−22,17)/ns |

**Negative**

- Inferior Frontal Gyrus | 47 | 0.4/0.0 | 3.9(−34,28,−4)/ns |
- Insula | 13 | 0.1/0.0 | 3.6(−31,22,−6)/ns |
3.4 Discussion

In this study we used GLM to generate activation maps for multiple fMRI conditions related to speech perception and comprehension, and jICA to decompose the activation maps into independent maps that share modulation profiles. This is similar to assuming a fixed hemodynamic response for each condition.

Figure 3.3: The effect of combining different contrasts on differentiation between histograms. The higher the values of the Renyi divergence, the better the discrimination between groups.

Fig. 3.7 shows the ROC curves and the corresponding detection reliability when the analysis is run with different numbers of ICs. It is easily seen that adding more ICs has little effect on the performance. Comparing the classification errors, there is really no major difference between different numbers of ICs.

3.4 Discussion
3.4. Discussion

Figure 3.4: Classification accuracy for different number of features used. Selecting only three significant features (p-value < 0.05), produced an average classification accuracy of around 75%.

subject and modeling the amplitude differences in modulation profiles [31]. The modulation profiles were used within a non-linear SVM framework to classify individuals. Our major findings are: (1) brain functional patterns of activation permit classification of individuals as younger or older; (2) combining these patterns improves the separability of joint sources and the accuracy of classification of individuals.

The joint source that differed the most between the two groups appeared to reflect activation differences at multiple levels of processing, including in left inferior frontal cortex for the contrast between coherent and anomalous sentences, in both temporal lobes for the contrast between anomalous sentences and SCN, and in primary auditory cortex bilaterally for the contrast of SCN vs. silence. This reflects the well-known hierarchy of speech processing in which low-level acoustic features are analyzed in auditory cortex,
3.4. Discussion

Figure 3.5: ROC curves and detection reliability resulting from four different numbers of features used. Perfect detection: $\rho = 1$; detection failure: $\rho = 0$. Performance of the classification was measured by repeatedly splitting the data into training and test sets and averaging classification performance on test data set. Splitting was done 200 times and the classification was repeated 5 times with randomized order of the subjects in the training dataset. ROC curves are computed using posterior probabilities of the SVM output.

Superior and middle temporal gyri are sensitive to processing of auditorily presented sentences, and left inferior frontal gyrus activity reflects higher-level linguistic (possibly semantic) processing [56, 55, 161, 142, 137].

The relation between the three maps of the joint sources (see Fig. 3.1) were investigated by looking back to the SPM contrast images and examining regions that contributed significantly in the joint source, i.e. computing the joint histogram. The divergence criterion derived from the joint histograms was used to measure the separability of the two groups based on the joint sources. This criterion confirmed that the fusion of contrasts improved sep-
3.4. Discussion

Figure 3.6: ROC curves and detection reliability for different contrasts. Perfect detection: $\rho = 1$; detection failure: $\rho = 0$. All of the contrasts (except that of coherent vs anomalous sentences) show high detection accuracy.

Results demonstrate that individuals can be classified relatively accurately into young and older age groups by combining functional contrasts sensitive to the processing of noise vs. silence, anomalous sentences vs. noise, and coherent vs. anomalous sentences. Note that although the brain imaging data permit this classification, behavioral data did not: the ability to report words from the anomalous and coherent sentences was matched between young and older listeners for the contrasts examined. Fig. 3.4 shows that using only three coefficients of the mixing matrix, a classification accuracy of around 75% can be obtained, albeit with a high standard deviation. The high standard deviation on the classification performance might be due to the small number of datasets, or because our method of controlling for hearing ability based on the behavioral performance did not work as well as
we had hoped.

Although using all three contrasts resulted in the best detection reliability, i.e. highest area under ROC curve, the contrast of anomalous sentences vs. unintelligible noise had the most impact on separability of the groups, with a detection reliability of around 60% (AUC of 80%) by itself. This may be because, in order to match intelligibility, older adults heard sentence materials at more advantageous SNRs, and the acoustic differences between these more positive SNRs and those experienced by the younger listeners may be reflected in different patterns of activity in auditory regions in the two groups. The analysis did not appear sensitive to the number of independent components (8, 12, or 16) included. Also, the number of features, as long as it was three or more, had relatively little impact on classification accuracy.

This was simply a validation study to demonstrate that information
3.4. Discussion

across multiple functional contrasts can be usefully combined for classification. Although here we differentiate young and older people, using age as an observable "gold standard" way to discriminate groups, we anticipate that this method will be useful to aid in classification of individuals to clinical groups using objective, quantitative, criteria. In chapter 7, we apply this method in order to classify individuals with depression.

In summary, using the joint ICA method together with an SVM classification algorithm, we have demonstrated that cognitive patterns can be used to classify individuals in the absence of behavioral differences. Feasibility of the proposed framework is shown by demonstrating that functional activity maps can be used to classify subjects accurately. The best combination of contrasts and optimal components are identified. We showed that by combining three different functional contrasts, revealing three different patterns of brain activity, the overall performance of the classification improves.
Chapter 4

Joint Sparse Representation Analysis

4.1 Introduction

A recent study by Daubechies et al. showed that the theoretical assumption of independence of the patterns extracted by ICA algorithms is not guaranteed in practice [51]. Furthermore, there is no physical reason for the spatial samples to correspond to different activity patterns with statistically independent distributions [125]. It has been shown that, by using ICA algorithms, patterns are separated on the basis of spatial sparsity rather than independence [51, 195, 114]; therefore, mathematical properties of brain data other than independence should be used.

Unlike Independent Component Analysis based methods, sparse representation methods do not require independence of the brain activation patterns. Although a recent study by Calhoun et al. claimed that ICA algorithms used for fMRI analysis select for independence rather than sparsity [32], several studies have shown superiority of sparse representation methods over the ICA-based methods [164, 103, 64]. Roussos et al. employed wavelets and sparsity-inducing adaptive priors to construct a structured generative latent-variable model, and showed their proposed algorithm outperforms ICA in benchmark datasets. Lee et al. used K-Singular Value Decomposition (K-SVD; [1]) to estimate design matrices within a sparse general linear model (GLM) framework [103]. Eavani et al. utilized K-SVD to identify distinct functional sub-networks of resting-state fMRI data [64]. Kim et al. employed sparse prior regularization for temporally concatenated ICA to enhance the estimation of an individual’s spatial patterns and the temporal components of neuronal activation [98]. Wang et al. showed better detection sensitivity of fMRI signal using a sparse approximation coefficient.

0This chapter is adapted from [154]: Mahdi Ramezani, Kristopher Marble, Heather Trang, Ingrid Johnsrude, and Purang Abolmaesumi, Joint Sparse Representation of Brain Activity Patterns in Multi-task fMRI Data, IEEE Transactions on Medical Imaging, 2014.
prior to ICA decomposition [202]. These studies demonstrated the superiority of sparse representation methods over ICA-based algorithms.

4.2 Method

In multi-task fMRI, the functional contrasts are obtained from subjects’ fMRI responses to different stimuli and tasks, which produce multiple functional contrasts. It is reasonable to expect that these functional contrasts contain some shared information, because even distant brain regions are often connected [127]. Our goal is to examine brain activations across multiple tasks, collected on the same participants, in a unified analytic framework by modeling potential coupling among the data from different fMRI tasks (the features). Here, the features are contrast images generated from multiple tasks. Each feature is represented as a set of sparse, linearly mixed, joint brain maps, which is formally analogous to conventional sparse representation of signals.

Sparse representation methods assume a generative model $y = Dx$, where a dictionary, $D$, combines with the sparse coefficients $x$, to generate the observation $y$. The objective is to maximize the likelihood that observation has efficient, sparse representations in a redundant dictionary given by the matrix $D$ [191]. Formally, the goal of learning is to find the overcomplete dictionary $D^*$ such that

$$D^* = \arg \max_D \left[ \log \hat{P}(y|x) \right] = \arg \max_D \left[ \log \int P(y|x, D)P(x)dx \right] \quad (4.1)$$

If we were to run the sparse representation on data from each fMRI task separately, we would maximize the likelihood functions in separate sparse representation analysis as follows:

$$D_1^* = \arg \max_D \left[ \log \int P(y_1|x_1, D_1)P(x_2)dx_1 \right],$$

$$D_2^* = \arg \max_D \left[ \log \int P(y_2|x_2, D_2)P(x_2)dx_2 \right],$$

$$\vdots$$

$$D_M^* = \arg \max_D \left[ \log \int P(y_M|x_M, D_M)P(x_M)dx_M \right].$$

(4.2)

This would result in $M$ set of dictionaries $[D_1^*, D_2^*, \ldots, D_M^*]$ and sparse coefficients $[x_1^*, x_2^*, \ldots, x_M^*]$. To obtain the joint relation of the results we
4.2. Method

would need to combine the sparse coefficients. However, if we utilize a fusion approach, the dictionaries can be learned such that they efficiently describe the content of the fMRI contrasts and simultaneously allow us to capture the correlation among multi-task contrasts. The learning objective is to maximize the joint likelihood that all contrasts are well represented by a dictionary $D$.

$$D^* = \arg \max_D \left[ \log \int P(y^1, y^2, \ldots, y^M|x, D)P(x)dx \right] \quad (4.3)$$

where $y^j$ is the functional contrast obtained from subject’s fMRI responses to task $j$, and $D$ is the joint dictionary containing joint sources.

The flowchart of the analysis, including the preprocessing, feature selection, joint sparse representation analysis, and visualization of the output sources is displayed in Fig. [4.1]. Preprocessing steps include realignment of the functional images, coregistration of the functional images to structural MRI data of subjects, and normalization of each subject’s structural MRI data to a template, and using the deformation parameters to normalize the functional images. To select features, single-subject general linear models are created by coding the condition to which each scan of the fMRI session belonged. As a result, functional contrast images (the features) are generated for each subject that represent the brain activation patterns related to the specific task. The functional contrast images of each subject are normalized to have the same average sum of squares. Then, subjects’ functional contrasts are stacked together and a joint activation pattern (joint feature) is created. Observation matrix, $Y = [y_1, y_2, \ldots, y_N] \in \mathbb{R}^{V \times N}$, is created by joint features of all subjects, where $y_i = [y_{i1}^1, y_{i2}^2, \ldots, y_{iM}^M]^T$ is the vector containing functional contrasts of subject $i$ stacking together.

In jSRA, we assume a generative model $Y = DX$. In this model, $Y$ is the observation matrix, $X = [x_1, x_2, \ldots, x_N] \in \mathbb{R}^{M \times N}$ is the sparse modulation matrix, and $D = [d_1, d_2, \ldots, d_N] \in \mathbb{R}^{V \times M}$ is the dictionary containing $M$ signal atoms, where $d_i = [d_{i1}^T, d_{i2}^T, \ldots, d_{iM}^M] \in \mathbb{R}^{M \times N}$ is the vector containing functional maps of the dictionary atom $i$. Each of these atoms represents the joint activation pattern of brain extracted from the subjects’ functional contrasts. $V$, $N$, and $M$ are the number of total voxels, subjects and brain patterns, respectively.
4.2. Method

Figure 4.1: Flowchart of the analysis, including preprocessing, feature selection, joint sparse representation analysis, and visualizing the output sources.

Fig. 4.2 shows the schematic of the joint sparse representation analysis. Assuming that maps of joint sources share the sparse modulation matrix, the method can represent a large set of functional contrasts as a sparse linear combination of a small set of joint basis patterns. This method does
4.2. Method

Prior to the decomposition, the appropriate number of brain basis patterns (dictionary size) should be selected. The best way to estimate this number of brain patterns, which may affect the results of the sparse representation analysis, is not clear [64]. Information theoretic techniques have been shown to be useful for selecting the number of brain basis patterns [104]. However, these techniques may not converge because of the heterogeneity in localization and size of activations. Therefore, in this study we perform the analysis using different values for the number of brain patterns, and compare the results with each other. In the jSRA method, we use the K-Singular Value Decomposition (K-SVD) method to decompose the multiple fMRI contrast images into joint sources. K-Singular Value Decomposition (KSVD) method [1] is one dictionary-learning algorithm for sparse signal representations, which decomposes the input matrix, \( \mathbf{Y} \), into a linear combination of dictionary elements, \( \mathbf{D} \), using the fewest number of non-zero

\[
\mathbf{Y} = \mathbf{D} \mathbf{X} + \mathbf{E}
\]

Figure 4.2: Schematic of the joint Source Representation Analysis (jSRA) method.
4.3 Simulations

coefficients. In other words, it solves the following minimization problem:

$$\min_{D, x}\{\|Y - DX\|_F^2\} \text{ subject to } \forall i, \|x_i\|_0 \leq T_0$$  \hspace{1cm} (4.4)

In this equation, $\|\cdot\|_F$ is the Frobenius matrix norm and $T_0$ is the number of non-zero elements in each linear combination. To find the sparsest representation of input contrasts, KSVD iteratively updates the vectors $x_i$ and each column of the dictionary in two steps [1]. Assuming a fixed dictionary, in the first step, the following minimization problem is solved using the Orthogonal Matching Pursuit (OMP) algorithm [141]:

$$\min_{x_i}\{\|y_i - Dx_i\|_F^2\} \text{ subject to } \forall i, \|x_i\|_0 \leq T_0, \; i = 1, 2, \ldots, N.$$ \hspace{1cm} (4.5)

In the second step, for each column of the dictionary ($d_k$), the representation error ($E_k$) is computed and using SVD decomposition the updated dictionary column ($\hat{d}_k$) is obtained:

$$E_k = Y - \sum_{j \neq k} d_j x^j, \; E_k = U \Delta V^T, \; k = 1, 2, \ldots, K.$$ \hspace{1cm} (4.6)

where $x^j$ is the $j$th row in $X$. The first column of $U$ is chosen as $\hat{d}_k$ and the first column of $V$ multiplied by $\Delta(1, 1)$ is chosen as the updated $x^k$. These steps are run for a finite number of iterations (See [1] for more details).

The K-SVD procedure generates a set of joint sources and associated sparse coefficients. These low-dimensional coefficients model the modulation of each subject’s functional contrast by a joint source. Thus, in group analysis of fMRI data, these coefficients can be used as a criterion for capturing group differences. To investigate whether the groups are separable by different weightings of the joint sources, unpaired two-sample t-tests with unequal variance (heteroscedastic) on the mixing coefficients are performed. For visualization, the corresponding joint sources are converted to z-values.

4.3 Simulations

4.3.1 Data Generation

The simulated fMRI data were generated by taking experimental fMRI data acquired using sparse imaging technique during a task paradigm, and (a) randomizing the time-course of each voxel to remove the intrinsic task-related activations, (b) adding pre-defined spatial patterns of functional activation
with varying degrees of spatial overlap size across subjects, and (c) adding corresponding task-related hemodynamic response functions (HRFs) associated with the neuronal activation.

A total of 20 simulated datasets that represent two groups of subjects, each with 10 datasets, were generated. For each subject multiple datasets representing multiple tasks were created. To make the simulations more realistic, we have used the experimental fMRI data of a subject, which was acquired during a speech comprehension task (see Section 2.1). The data were acquired across four blocks of trials, each 81 trials long. Each block of data was used to generate a simulated dataset that represents fMRI data obtained during a single task. As each task can activate multiple regions, which may or may not be similar to those activated by other tasks, multiple regions of interest (ROI) were defined. Five regions of interest (ROI) including Broca’s area, left hand motor function, medial frontal cortex, right dorsolateral prefrontal cortex, and Wernicke’s area were selected from the non-overlapping functionally connected brain regions defined by the Pittsburgh Brain Connectivity competition (PBC 2009). The PBC functional maps are cortical areas that are associated with cognitive control networks [42], language function (Wernicke’s and Broca’s area [17]), motor function (hand, foot, and tongue [198]), retinal field maps [201], and auditory responsive cortex [71]. Fig. 4.3 shows the regions of interest.

For each group, we defined four tasks that activate similar ROIs in subjects within each group, and similar and/or different ROIs between the two groups of subjects. Table 4.1 shows the ROIs used for each group for each task. Task 1 is defined so that it includes two similar regions between the groups; task 2 and task 3 have one similar and one dissimilar regions; and task 4 has two dissimilar regions between the groups.

Since each activated region in each subject can be slightly different from other subjects within the same group, the ROIs were randomly dilated using a disk with variable radius of one to five voxels for each subject. Therefore, all subjects within each group have a common known area of activation.

For each ROI of each subject a synthetic activation signal was created using the SimTB MATLAB toolbox [68]. A block experimental design with 81 time samples, a repetition time of 9 sec, a block length of 6 sec, and inter-stimulus interval of 6 sec was defined. The activation signal was generated by linear convolution of the block design signal and a canonical HRF.

Rician noise [84] was added to real and imaginary parts of the activation signals and taking the square root relative to a specified contrast-to-noise ratio (CNR). CNR, which is the ratio of the standard deviation of the signal to the standard deviation of the noise [68], obtained from a uniform distribution
4.3. Simulations

Figure 4.3: ROIs representing the activated regions within the brain, defined using the brain functional connectivity maps provided by the Pittsburgh Brain Connectivity competition (PBC 2009). Blue, green, red, cyan and magenta colors show Broca, left hand, medial frontal cortex, right dorsolateral prefrontal cortex, and Wernicke.

of numbers between 0.5 and 1.5, for each of the signals.

Table 4.1: ROIs used for each group for each task.

<table>
<thead>
<tr>
<th>Task</th>
<th>group 1</th>
<th>group 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>broca, wernicke</td>
<td>broca, wernicke</td>
</tr>
<tr>
<td>2</td>
<td>medial frontal</td>
<td>left hand, right</td>
</tr>
<tr>
<td></td>
<td>cortex, right</td>
<td>dorsolateral</td>
</tr>
<tr>
<td></td>
<td>prefrontal cortex</td>
<td>prefrontal cortex</td>
</tr>
<tr>
<td>3</td>
<td>broca, left hand</td>
<td>left hand, wernicke</td>
</tr>
<tr>
<td>4</td>
<td>medial frontal</td>
<td>broca, right</td>
</tr>
<tr>
<td></td>
<td>cortex, wernicke</td>
<td>dorsolateral</td>
</tr>
<tr>
<td></td>
<td></td>
<td>prefrontal cortex</td>
</tr>
</tbody>
</table>

After realigning the data set using SPM8, the time course of each voxel was randomized to remove the intrinsic activations in the whole dataset. Then, for each subject, the activation signal of each ROI was added to the
4.3. Simulations

Figure 4.4: fMRI signal of a voxel before (a) and after randomization (b), the created synthetic activation signal (c), and simulated fMRI signal (d). A simulated fMRI signal is generated by adding the weighted activation signal to the randomized fMRI signal.

time courses of voxels in that ROI. The amplitude of the activation signal was chosen to be 3% of the mean amplitude of the time course of the selected voxel for the subjects in group 1, and to be 5% of the mean amplitude of the time course of the selected voxel for the subjects in group 2. Fig. 4.4 shows the fMRI signal of a voxel before (a) and after randomization (b), the synthetic activation signal (c), and simulated fMRI signal (d), which is created by adding the weighted activation signal to the randomized fMRI signal.

4.3.2 Experiments and Results

The simulated data were spatially smoothed using Gaussian kernel of 8 mm, and the contrast images related to the simulated conditions were created. In order to generate contrast images related to the simulated conditions, single-subject general linear models were created. Different combinations of contrasts generated from each task were used as input observations to the jSRA method. Besides the combination of all four tasks, two combinations of three tasks were generated, by removing the task that activated two similar
4.3. Simulations

regions between the groups, i.e. task 1, and by removing the task that activated two dissimilar regions between the two tasks, i.e. task 4. See Table 4.1 for regions in each task. Combinations of two tasks were generated as follows: Tasks 1 and 2 that do not have any overlap among the synthetic activation regions, i.e. broca and wernicke regions for task 1 and medial frontal cortex, right dorsolateral prefrontal cortex and left hand regions for task 2; Tasks 3 and 4 that have one activation map for each group being opposites of each other, i.e. the activation map in task 3 of group 1 is similar to the task 4 of group 2 and vice versa; Tasks 1 and 4 have one similar activation map for both groups, i.e. wernicke region in group 1 and broca region in group 2; Tasks 2 and 3 have one similar activation map for one of the groups, i.e. left hand region in group 2.

Using the jSRA method, the joint sources for each combination were obtained. These joint sources are estimations of the true joint sources, i.e. activation maps that were used to generate the observations. We compared the estimated and true joint sources by calculating the False Positive (FP), False Negative (FN), True Positive (TP) and True Negative (TN) values. TP shows the number of correctly identified active voxels, TN shows the number of correctly identified non-active voxels, FP shows the number of non-active voxels identified as active voxels, and FN shows the number of active voxels identified as non-active voxels, in the estimated joint sources. Furthermore, the similarity indices such as Jaccard, precision, and sensitivity were calculated, as defined below:

\[
Jaccard = \frac{TP}{TP + FP + FN}, \quad (4.7a)
\]

\[
Precision = \frac{TP}{TP + FP}, \quad (4.7b)
\]

\[
Sensitivity = \frac{TP}{TP + FN}. \quad (4.7c)
\]

As there are many more non-active voxels in the brain than active voxels, TN is a much bigger number than TP, FP, and FN. The specificity and accuracy of the results will be near 99%, and will not differ between the results of the jSRA and jICA. Accordingly we do not report these indices. Similar to the jSRA method, the jICA method [31] was applied to the same combinations of contrasts and similarity coefficients were calculated.

Using the estimated maps of the joint sources for the jSRA and the jICA methods, the Jaccard index, precision, and sensitivity were calculated. Table 4.2 shows the results. As it can be seen, the jSRA method is superior to the jICA method. The best performance occurs for the combination of
Table 4.2: Similarity coefficients showing the result of jICA and jSRA methods for different combinations of tasks. jS1 and jS2 stand for joint source 1 and 2.

<table>
<thead>
<tr>
<th>Tasks</th>
<th>Method</th>
<th>Jaccard</th>
<th>Precision</th>
<th>Sensitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>jS1</td>
<td>jS2</td>
<td>jS1</td>
</tr>
<tr>
<td>T1+T4</td>
<td>jICA</td>
<td>49.2</td>
<td>54.1</td>
<td>62.7</td>
</tr>
<tr>
<td></td>
<td>jSRA</td>
<td>52.9</td>
<td>63.2</td>
<td>68.5</td>
</tr>
<tr>
<td>T2+T3</td>
<td>jICA</td>
<td>53.8</td>
<td>39.3</td>
<td>82.5</td>
</tr>
<tr>
<td></td>
<td>jSRA</td>
<td>62.5</td>
<td>52.9</td>
<td>74.8</td>
</tr>
<tr>
<td>T1+T2</td>
<td>jICA</td>
<td>40.3</td>
<td>56.4</td>
<td>53.7</td>
</tr>
<tr>
<td></td>
<td>jSRA</td>
<td>47.1</td>
<td>59.2</td>
<td>59.9</td>
</tr>
<tr>
<td>T3+T4</td>
<td>jICA</td>
<td>46.4</td>
<td>67.4</td>
<td>60.6</td>
</tr>
<tr>
<td></td>
<td>jSRA</td>
<td>60.1</td>
<td>66.6</td>
<td>74.0</td>
</tr>
<tr>
<td>T1+T2+T3</td>
<td>jICA</td>
<td>37.5</td>
<td>58.2</td>
<td>50.1</td>
</tr>
<tr>
<td></td>
<td>jSRA</td>
<td>51.9</td>
<td>62.2</td>
<td>64.9</td>
</tr>
<tr>
<td>T2+T3+T4</td>
<td>jICA</td>
<td>26.6</td>
<td>54.9</td>
<td>36.2</td>
</tr>
<tr>
<td></td>
<td>jSRA</td>
<td>53.7</td>
<td>62.7</td>
<td>65.6</td>
</tr>
<tr>
<td>T1+T2+T3+T4</td>
<td>jICA</td>
<td>30.7</td>
<td>35.1</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>jSRA</td>
<td>52.6</td>
<td>43.3</td>
<td>66.3</td>
</tr>
</tbody>
</table>

tasks 3 and 4, where the average Jaccard index of 63.35% is obtained, and the worst performance is for the combination of all tasks, where an average Jaccard index of 47.96% is obtained.

As the number of combined tasks, which is related to the dimensionality of the input observations, increases, the values of the similarity indices between the estimated and true maps drops. However, the jSRA method is more robust to the increase in the dimensionality. In other words, the improvement of the jSRA method, in comparison to jICA, becomes more marked as the number of combined tasks increases. Fig. 6 (a, b, c) shows the average Jaccard, precision and sensitivity for combinations of two, three and four tasks.
4.4 Results

For visualization, the joint source maps were converted to empirical Z-values by subtracting the mean of the joint source from each value and then dividing each result by the standard deviation. Figs. 4.6, 4.7, and 4.8 show the statistical Z maps (|Z| > 2.25) generated for the most significant joint source (i.e., the one with the smallest p-value) obtained using the statistical t-test on the sparse/mixing coefficients for the jSRA/jICA methods. The three figures correspond to analyses with 4, 8 and 12 components, respectively. Figs. 4.6(a, e), 4.7(a,e), and 4.8(a,e) show the distribution of the sparse/mixing coefficients with the most significant difference between the

Figure 4.5: Average Jaccard (a), precision (b), and sensitivity (c) for combinations of two, three and four tasks.

4.4 Results

For visualization, the joint source maps were converted to empirical Z-values by subtracting the mean of the joint source from each value and then dividing each result by the standard deviation. Figs. 4.6, 4.7, and 4.8 show the statistical Z maps (|Z| > 2.25) generated for the most significant joint source (i.e., the one with the smallest p-value) obtained using the statistical t-test on the sparse/mixing coefficients for the jSRA/jICA methods. The three figures correspond to analyses with 4, 8 and 12 components, respectively. Figs. 4.6(a, e), 4.7(a,e), and 4.8(a,e) show the distribution of the sparse/mixing coefficients with the most significant difference between the
4.4. Results

Figure 4.6: Significant joint source for jSRA (a-d) and jICA (f-h) methods, using 4 sources. From left to right the figures show the distributions of the coefficients, and the corresponding maps related to high-, mid-, and low-level auditory processing. White ovals show the most expected activation regions: 1) left inferior frontal cortex for the high-level contrast; 2) temporal lobes, particularly in the left hemisphere, for the mid-level contrast; and 3) primary auditory cortex in both hemispheres for the low-level contrast.

two groups. As it can be seen, the absolute value of the modulation coefficients is higher for older compared to young adults. These coefficients provide a measure of functional connectivity [31] implying greater functional connectivity among different levels of auditory processing in older adults.

Fig.4.6e shows that jICA has failed to identify a joint source that is significantly different between the two groups, when the number of estimated independent components was four. We expected to see joint activation maps in the left inferior frontal cortex for the high-level contrast, in the temporal lobes, particularly in the left hemisphere, for the mid-level contrast, and in the primary auditory cortex in both hemispheres for the low-level contrast [186]. These regions are shown using white ovals in Figs. 4.6, 4.7, and 4.8. As it can be seen in these figures, both of the jICA and jSRA methods showed almost all of the regions using 4, 8, and 12 sources. However, the
4.5 Discussions

In this study, the jSRA method was used to identify patterns of brain activity (fMRI) evoked by multiple cognitive tasks within subjects. Our major findings are: (1) sparse representation analysis for multi-task fMRI data analysis better captures the activation maps compared to the jICA approach; (2) brain functional-activation patterns can be represented as sparse coefficients...
4.5. Discussions

Figure 4.8: Significant joint source for jSRA (a-d) and jICA (f-h) methods, using 12 sources. From left to right the figures show the distributions of the coefficients, and the corresponding maps related to high-, mid-, and low-level auditory processing. White ovals show the most expected activation regions: 1) left inferior frontal cortex for the high-level contrast; 2) temporal lobes, particularly in the left hemisphere, for the mid-level contrast; and 3) primary auditory cortex in both hemispheres for the low-level contrast.

that capture individual differences. We have shown these findings within two experiments using simulated fMRI data and experimental fMRI data from a speech comprehension task.

Simulation experiments showed the superiority of the jSRA method compared to the jICA method in terms of the similarity indices between the estimated and true activation maps. As the number of input tasks increases, the average value of the similarity indices, for both of the jICA and jSRA methods, become smaller. This is reasonable considering the fact that the number of observations, i.e. number of subjects, does not change, but the dimensionality of the input observations is multiplied by a factor of 1.5 or two. However, jICA was more sensitive to the increase in the dimensionality of the input observations. The change in the jSRA results related to the dimension increase was less than that for the jICA method. Therefore the relative improvement in sensitivity and specificity, for jSRA compared to
4.5. Discussions

jICA, also becomes more marked, as the number of combined tasks increases. Looking more closely at the results of the analysis of simulated data, we can see that, among the combinations of two tasks, combination of tasks one and two had the worst performance. In this combination of tasks, there was no overlap among the simulated activated regions between the two tasks. Besides, three of these four regions were similar between the two groups, i.e. broca, wernicke and right dorsolateral prefrontal cortex regions. Therefore, it was difficult to estimate the true sources and capturing the group differences regardless of method.

Combination of tasks three and four resulted in the greatest differentiation between the groups, with the activation patterns for the two groups being opposites of each other. Group differences were captured not only by the amplitude differences of the modulation profiles, but also through the complementarity of activated regions in the two groups: the activation map for task 3 of group 1 partially matched that for task 4 of group 2, and vice versa.

A concern about the simulations is losing temporally correlated noise structure (e.g., physiological noise) by randomizing the time-course of the fMRI signals. However, since we simulated sparse imaging data with 9 sec between scans, it is reasonable to assume that successive scans are largely uncorrelated.

Results of the real experimental data showed that the activation maps that were captured within the maps of the joint sources, which were significantly different between the two groups, reflected the well-known hierarchy of speech processing: Auditory cortex analyzes the low-level acoustic features; superior and middle temporal gyri are sensitive to processing of auditorily presented sentences; and left inferior frontal gyrus activity reflects semantic processing [56, 55, 161, 142, 137]. Sensitivity of the analysis to the number of brain basis patterns was examined, by choosing 4, 8, and 12 numbers of patterns. We showed that unlike the jICA, the proposed jSRA method is not very sensitive to the chosen number of components. Although the difference between jSRA and jICA appears to be reduced with more components, using more components introduces two problems when dealing with group fMRI data acquired from multiple subjects: 1) selecting the components of interest, and 2) split activation patterns within a region of interest.

An additional advantage of jSRA method is that the modulation profiles are substantially compact and sparse. These coefficients can be used for reducing the dimensionality of the fMRI data, and may allow for a more reliable classification.

In summary, we have shown that a joint sparse representation analysis
can effectively identify the common and unique information among different levels of brain cognitive patterns in multi-task fMRI data within different groups. To demonstrate the potential of the proposed framework, analyses of simulated fMRI data were performed followed by analyses of experimental fMRI data from normal subjects performing speech comprehension tasks. Simulations showed the superiority of the proposed method to the state of the art method (jICA) for multi-task fMRI data analysis. Results on the experimental fMRI data also demonstrate that the jSRA method can better capture the brain functional activation patterns, and therefore the differences, between two groups.
Chapter 5

Reliability Analysis and Visualization

5.1 Introduction

Sparse representation methods have gained popularity for analysis of brain data sets including EEG [178, 10, 86], MRI [181, 78], and fMRI [107, 106, 164, 149], in recent years. In functional MRI studies, sparse representation methods have been used for multivariate pattern analysis (MVPA) [79, 106, 181, 166, 33], statistical parametric mapping for detection of task related activation [103], classification of individuals [149, 150], and identification of brain resting-state networks [64]. Sparse representation analysis decomposes the input observations into a linear combination of dictionary atoms, using the fewest number of non-zero coefficients.

Despite the great advantage of sparse representation methods, there are significant drawbacks associated with these methods. First, the results of most sparse-representation algorithms may be somewhat different in multiple runs of the algorithm, depending on the initialization of dictionary atoms. It has been shown that the exact determination of sparsest representation is an NP-hard problem [54]. Therefore, approximate solutions have been proposed for instead. Most sparse representation algorithms solve a minimization problem iteratively and guarantee convergence to at least a local minimum solution. Depending on the initialization and parameters value, the algorithms may find different local minima. Therefore, an algorithmic reliability analysis of estimated sparse dictionary-atoms, hereafter named components, is needed. Algorithmic reliability refers to the fact of finding the point that globally minimizes (maximizes) the objective function. Moreover, robustness of the sparse representation analysis with respect to the parameters, i.e. dictionary size and sparsity level, should be investi-

\[\text{This chapter is adapted from the following submission [155]: Mahdi Ramezani, Saman Nouranian, Ingrid Johnsrude, and Purang Abolmaesumi, Reliability Analysis and Visualization of Sparse Representation Methods for Neuroimaging Data, submitted, 2014.}\]
gated prior to interpretation of the results. Algorithmic reliability should not be mistaken by the statistical reliability of sparse representation analysis which is a necessary step for interpretation of the results of any statistical method. Such analysis can be accomplished by bootstrapping (resampling) or variable selection (subsampling) algorithms [65, 8, 126]. Within these resampling/subsampling methods the data sample is randomly changed by simulating the sampling process, and the algorithm is run multiple times with the resampled/subsampled data. The spread of the obtained components is used to assess the statistical reliability of the original estimated component [91].

Second, the dictionary atoms (components) are not ordered and are randomly permuted. This introduces two problems, when dealing with group fMRI data acquired from multiple subjects, where the output components may be split activation patterns within a region of interest (ROI) or merged activation patterns of multiple ROIs [98, 91, 124]: 1) selecting the components of interest, and 2) choosing the optimum number of components. The first problem, is more important where no prior information of the spatial location of the activation patterns is available. The second problem arises from the fact that fMRI signal consists of multiple components such as hemodynamic changes due to neural activations which can be task related or non-task related, motion and MRI scanner artifacts, and cardiac and respiratory pulsations. If the number of output components is selected less than the actual number of components, the output component may appear as merged activation patterns; and if the number of output component is selected far more than the actual number of components, the output components may be split activation patterns.

A handful of studies have tried to overcome these issues. Eavani et al. used a Hausdorff distance metric to compare the dictionaries obtained by running the algorithm for different values of dictionary size and sparsity level, and ranked the estimated components [64]. Bilwaj et al. used $L^2$ norm of the coefficients of a corresponding component to rank the components [78]. In our previous work, we ranked the components based on the ability of the estimated components to classify individuals [149]. However, these studies did not address the effect of variation in components in multiple runs in these papers, nor did they investigate the robustness of the method with respect to the parameters (i.e. dictionary size and sparsity level).

To circumvent these issues of component splitting and merging as well as random permutation and to facilitate use of the sparse representation method for group fMRI data, we present a method for assessing the reliability of the estimated components as well as for visual inspection of components in 2D
5.1. Introduction

space, using a non-linear mapping. The method is based on the work by Himberg et al. [91], ICASSO, which was proposed for stability estimation of independent components. We run the sparse-representation algorithm multiple times with different initializations and cluster the estimated components based on the similarity of the components as represented by a correlation coefficient. We further visualize the similarities using a nonlinear 2D projection: t-SNE [59]. T-SNE visualizes high-dimensional data by assigning each data point a location in two-dimensional map. Visualization allows further interpretation of the clustering results, including individual estimates within cluster, and the relations between clusters. Using t-SNE visualization the similarity of the components will be clearly presented. Unlike the ICASSO method [91], the t-SNE method maintains the local and global structure of the data in a single map, and keeps the quality of the visualization when the number of components becomes large [59].

In this chapter, we examine the reliability with which sparse analyses can reliably identify activation foci in multi-subject functional imaging studies. We first examine this in simulated data, and then extend our analysis to real data. We use simulated fMRI data sets, four groups of subjects with 10 subjects per group and two activation foci in all subjects of each group, to quantitatively evaluate the results. Simulated fMRI data were generated by taking experimental fMRI data (described in Material section), and (a) randomizing the time-course of each voxel to remove the intrinsic condition-related activations, (b) adding pre-defined spatial patterns of functional activation with varying degrees of spatial overlap size across subjects, and (c) adding corresponding task-related hemodynamic response functions (HRFs) associated with the neuronal activation. We perform reliability analysis and visualization of the estimated fMRI components. We compare the results obtained using sparse analysis to those obtained using ICA, since this is another popular approach. We furthermore show that using the t-SNE visualization of the similarities between activation foci in multiple runs of algorithm, correct number of sources that were introduced in the simulated data could be detected. The assumption in the simulated data that all activations are found in all subjects may not be valid in real fMRI data, moreover, the intersubject variability in the simulated data may not be enough, therefore we report an analysis on experimental fMRI data acquired from sixteen subjects while performing tasks related to speech comprehension. We demonstrate how the proposed reliability assessment works if activation foci are not found in every subject, and if there is a realistic amount of intersubject variability. We show that the proposed approach helps in identifying the parameters such as dictionary size and sparsity level that result in reliable sparse sources.
5.2 Method

The proposed approach consists of four steps: a) Preprocessing of the fMRI data; b) setting the parameters of the sparse representation analysis and running it multiple times, with different dictionaries, using the selected parameters; c) clustering the components based on their similarities using the hierarchical clustering algorithm; d) visualizing the clusters using the t-SNE approach. The flowchart of the analysis is displayed in Fig. 5.1.

Figure 5.1: Flowchart of the analysis, including preprocessing, sparse representation analysis, reliability analysis, and visualization.
5.2. Method

5.2.1 Preprocessing

Preprocessing steps included realignment of the functional images, coregistration of the functional images to structural MRI data of subjects, and normalization of each subject’s structural MRI data to a template, and using the deformation parameters to normalize the functional images. Single-subject General Linear Models are created by coding the condition to which each scan of the fMRI session belonged. As a result, functional contrast images (the features) are generated for each subject that represent the brain activation patterns related to the specific task comparison. The functional contrast image of each subject are normalized to have the same average sum of squares. Observation matrix, \( Y = [y_1, y_2, \ldots, y_N] \in \mathbb{R}^{V \times N} \), is created by features of all subjects, where \( y_i \) is the vector containing functional contrast, \( V \) and \( N \) are the number of variables and subjects, respectively.

5.2.2 Sparse Representation Analysis

The goal of sparse representation approaches, when they are used as second level (group-level) analysis of fMRI data, is to represent functional contrast images as a set of sparse and linearly mixed brain maps. Sparse representation methods assume a generative model \( Y = DX \) [191]. In this model, \( Y \) is the observation matrix, \( X = [x_1, x_2, \ldots, x_N] \in \mathbb{R}^{K \times N} \) is the sparse modulation matrix, and \( D = [d_1, d_2, \ldots, d_N] \in \mathbb{R}^{V \times K} \) is the dictionary containing \( K \) signal atoms (estimated components) representing the activation pattern of brain extracted from the subjects’ functional contrasts. The objective is to maximize the likelihood that observation has efficient, sparse representations in a redundant dictionary given by the matrix \( D \) [191]. Here, we use the KSVD method, as a well-known sparse representation algorithm [1], to decompose the input matrix, \( Y \), into a linear combination of dictionary elements, \( D \), using the fewest number of non-zero coefficients. The number of iterations, components, and non-zero elements in each linear combination is set prior to running K-SVD. To initially estimate the number of components (dictionary elements) in the K-SVD analysis, we used the Minimum Description Length (MDL) criterion [104], which is an information theoretic technique for model order selection.

5.2.3 Reliability Analysis

The K-SVD algorithm is run \( M \) times on the data matrix \( Y \), with different initializations of the dictionary, \( D \), and the estimates of the demixing matrices, which are pseudo inverse of sparse modulation matrices, are collected
5.2. Method

into a single matrix \( W = [W_1^T, W_2^T, \ldots, W_M^T] \), where \( W_i^T = X_i^{-1} \).

Similarity of the estimated components is measured by the absolute value of their mutual correlation coefficients \( \rho_{ij}, i, j = 1, 2, \ldots, K \). Assuming the generative model \( Y = DX \), and \( YW^T = D \), the correlation coefficients are computed by:

\[
P = D^TD = WY^TYW^T = W(Y^TY)W^T = WCW^T, \tag{5.1}
\]

where \( C \) is the covariance matrix of the observation matrix, \( Y \) [91]. Hierarchical clustering algorithm [83] is used to cluster the estimated components using the dissimilarity measure defined as \( 1 - |\rho_{ij}| \). Using the hierarchical clustering algorithm the dendrogram is generated and partition of \( L \) clusters (which is set equal to the number of components) is generated by the average-linkage criterion. The number of points in each cluster is computed. These numbers, which shows the number of times each component is estimated by the sparse analysis, together with the similarities between estimated clusters, are used within the visualization step to select reliable components.

5.2.4 Visualization using t-SNE

The t-Distributed Stochastic Neighbor Embedding (t-SNE) algorithm [59] is used to visualize the clustering results in a two-dimensional map. t-SNE converts the Euclidean distance between the data points into conditional probabilities that represent similarities. Similarity in the high-dimensional input space between points \( j \) and \( i \) is the conditional probability \( p_{ij} \), which shows the probability that \( i \) would pick \( j \) as its neighbour to the probability density under a Student-t distribution centred at \( i \). A similar conditional probability can be computed in the low-dimensional (2D) space denoted by \( q_{ij} \). In a correct non-linear 2D projection map of the high-dimensional space to the 2D space, the two conditional probabilities, \( p_{ij} \) and \( q_{ij} \), will be equal. Therefore, minimizing the Kullback-Leibler divergence between the conditional probabilities over all input points will yield a correct 2D representation of the input space.

\[
Cost = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{q_{ij}} + \sum_i \sum_j q_{ij} \log \frac{q_{ij}}{p_{ij}}. \tag{5.2}
\]

t-SNE uses a symmetric KL-divergence criterion and minimizes it using a gradient descent method.

Each component is shown as a point in 2D space and the similarity of the components are represented by connected lines whose thickness denote the
similarities between them. A convex hull bounds the points belonging to the same cluster. Compact and isolated clusters represent reliable estimation of the components.

In order to evaluate the quality of the visualization, the 2D projection of t-SNE is compared to other projection techniques such as Curvilinear Component Analysis (CCA) [57] and Multidimensional Scaling (MDS) [190] which are typically used for 2D visualization. The quality of the projection is assessed based on the quality of the obtained clusters in two different ways. First, a hierarchical clustering algorithm is used to cluster the 2D projection of components. The quality of the clustering is evaluated using the cophenetic correlation coefficient which shows how well the dendrogram of the hierarchical clustering preserves the distances between the projected components [179]. The closer the value of this coefficient is to 1, the more accurately the clustering solution reflects the 2D projected components. Second, the trustworthiness index is used to compare the quality of the projection [197]. A projection onto a lower dimension is trustworthy if the set of $m$ closest neighbors of a point on the lower dimension are close by in the original space.

5.3 Simulations

5.3.1 Data Generation

A total of 40 simulated datasets that represent four groups of subjects, each with 10 datasets, were generated. The simulated data is generated similar to the 4.3.1. we have used the experimental fMRI data of a subject, which was acquired during a speech comprehension task (see Section 2.1). The data was acquired across 81 trials.

As each simulated subject activates multiple regions of activation, which may or may not be similar to those activated by others, multiple regions of interest (ROI) were defined. These ROIs were selected from the non-overlapping functionally connected brain regions shown in Fig. 4.3. For each group, two similar ROIs in subjects within each group, and similar and/or different ROIs between the other groups of subjects, were selected. Table 5.1 shows the ROIs used for each group. Since each activated region in each subject can be slightly different from other subjects within the same group, ROIs were randomly dilated using a disk with variable radius of one to five voxels for each subject. Therefore, all subjects within each group have a common known area of activation.

For each ROI of each subject a synthetic activation signal was created
5.3. Simulations

Table 5.1: ROIs used for each group to create the simulated data. Each group has two ROIs which are similar and/or dissimilar to other groups.

<table>
<thead>
<tr>
<th>Group</th>
<th>Activated Regions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Broca’s area</td>
</tr>
<tr>
<td></td>
<td>Wernicke’s area</td>
</tr>
<tr>
<td>2</td>
<td>medial frontal cortex</td>
</tr>
<tr>
<td></td>
<td>right dorsolateral prefrontal cortex</td>
</tr>
<tr>
<td>3</td>
<td>Broca’s area</td>
</tr>
<tr>
<td></td>
<td>left-hand motor cortex</td>
</tr>
<tr>
<td>4</td>
<td>medial frontal cortex</td>
</tr>
<tr>
<td></td>
<td>Wernicke’s area</td>
</tr>
</tbody>
</table>

as described in [4.3.1]. After realigning the data set using SPM8, the time course of each voxel was randomized to remove the intrinsic activations in the whole dataset. Then, for each subject, the activation signal of each ROI was added to the time courses of voxels in that ROI. The amplitude of the activation signal was chosen to be 3% of the mean amplitude of the time course of the selected voxels.

5.3.2 Experiments and Results

Simulated data were spatially smoothed using a Gaussian kernel of 8 mm, and the contrast images related to the simulated conditions were created. In the simulated dataset, four groups of subjects each with specific spatial patterns that were not the same between the groups were created. Hence, the correct number of dictionary size and sparsity level are four and one, respectively. However, in real cases these numbers are not known. Therefore, here we tried two different values for the dictionary size and sparsity level and investigated the results of the analysis. For various parameters setting, the dendrogram created by the hierarchical clustering algorithm and the number of KSVD components in the estimated clusters were plotted. The cluster of components was visualized in 2D space. Furthermore, the estimated components were compared to the true components and true positive rate (TPR) and false positive rate (FPR) of the estimated active voxels in the spatial patterns were calculated. Receiver operating characteristic (ROC) curves were computed, and the area under the curves were measured to quantitatively evaluate the performance of the estimated spatial patterns. To have a better understanding of the quality of the estimated components, results were compared to the conventional ICA analysis.

Fig. 5.2 shows the number of components in each cluster, the dendrogram of the hierarchical clustering algorithm, similarities between the estimated clusters, and 2D visualization of the clusters, for two different parameters.
5.3. Simulations

setting. The first row shows the result for dictionary size of 4 and sparsity level 1, and the second row shows the results for dictionary size 8 and sparsity level 4. The first column shows the number of times each component appeared in a cluster. The second column is the dendrogram of the hierarchical clustering algorithm obtained by the linkage strategy criterion. The third column shows the similarities between estimated clusters, arranged according to the dendrogram. The third row shows the 2D projection of the clusters using t-SNE approach for both parameter settings. Single run estimation of components, the centrottype of the clusters and the similarities between the components are shown in the third row figures. Convex hulls are generated around the estimated clusters. Compact and isolated clusters represent reliable estimation of the components. Fig. 5.2g shows that there are four compact and isolated clusters of components, which have appeared equally in multiple runs of the algorithm (Fig. 5.2a). Therefore, if the size of dictionary and sparsity level are correctly set, compact and isolated clusters will be generated. Fig. 5.2h shows that if these parameters are not correctly selected, the clusters are not compact and isolated. However, still separate cluster of components (in our experiment 3-4 clusters) are visible within the data, which can help with correctly identifying the parameters prior to re-running the algorithm.

To better investigate the results, the activation maps obtained using the two parameter settings is shown in Fig. 5.5. The first row shows the result for dictionary size of 4 and sparsity level 1, and the second row shows the results for dictionary size 8 and sparsity level 4. Only the first four components which appeared more than the other components are shown for the dictionary size of 8. In this figure, column 1 to 4 show activations related to group 1 to 4, respectively. The component number which is written below each activation map shows the component label in Fig. 5.2. These components’ numbers correspond to the large and isolated clusters for dictionary size 8 (Fig. 5.2h). Therefore, we could correctly identify the informative components by looking at the visualization, and selecting the largest compact clusters. In other words, the meaningful components are the ones with isolated and compact clusters which have appeared relatively more in multiple runs of the algorithm. This means that we can select meaningful components, even when the parameters are not initially correctly selected.

Fig. 5.3 shows the trustworthiness index as a function of neighborhood size for the CCA, MDS and t-SNE methods. According to this figure, the t-SNE method produces more trustworthy projections than CCA and MDS methods. The maximum trustworthiness occurs at around the neighborhood size of 20 which is equal to the number of times that the algorithm was run.
5.3. Simulations

Figure 5.2: Reliability and visualization analysis of the simulated data for two different sets of parameters (first and second row). (a, d) Number of times each component appeared in a cluster. (b, e) Dendrogram of the hierarchical clustering algorithm. (c, f) Similarities between estimated clusters, arranged according to the dendrogram. (g, h) 2D projection of the clusters.
5.3. Simulations

Figure 5.3: Trustworthiness index as a function of neighborhood size.

Figure 5.4: Distribution of cophenet correlation coefficient in multiple runs of the visualization and clustering. The central red mark is the median, the edges of the blue box are the 25th and 75th percentiles, and the whiskers show the extreme values of the volumes.

Fig. 5.4 shows the distribution of cophenet correlation coefficient in multiple runs of the visualization and clustering for the CCA, MDS and t-SNE methods. This figure shows that the t-SNE projection can significantly better preserve the pairwise distances between the projected components compared to CCA and MDS.

Table 5.2 shows the normalized AUC for the four estimated components using sparse analysis. As it can be seen the sparse analysis has successfully identified the activation maps related to each group. To have a better un-
5.4 Experiments and Results

For real fMRI experiments, we use the experimental fMRI data described in [2.1]. We use the conditions of listening to “anomalous” sentences, and the condition of silence. For each subject functional contrast image comparing

Independent Components (ICs) were found using the Infomax algorithm [12].
5.5 Discussions

Table 5.2: Normalized AUC for the four estimated components using the sparse analysis and ICA methods. C1 to C4 represent the true components within the simulated dataset.

<table>
<thead>
<tr>
<th>dict. size</th>
<th>sparsity level</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
</tr>
</thead>
<tbody>
<tr>
<td>sparse analysis</td>
<td>4</td>
<td>1</td>
<td>0.49</td>
<td>0.58</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>3</td>
<td>0.40</td>
<td>0.71</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>4</td>
<td>0.53</td>
<td>0.62</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>4</td>
<td>0.50</td>
<td>0.58</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>1</td>
<td>0.38</td>
<td>0.61</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>4</td>
<td>0.48</td>
<td>0.57</td>
<td>0.93</td>
</tr>
<tr>
<td>number of components</td>
<td>3</td>
<td>Failed</td>
<td>0.72</td>
<td>0.90</td>
<td>Failed</td>
</tr>
<tr>
<td>ICA</td>
<td>4</td>
<td>0.58</td>
<td>0.86</td>
<td>0.93</td>
<td>Failed</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.65</td>
<td>0.84</td>
<td>0.93</td>
<td>Failed</td>
</tr>
</tbody>
</table>

Fig. 5.6 shows the results of experimental fMRI data analysis. The number of times each component appeared in a cluster, the dendrogram of the hierarchical clustering algorithm, the similarities between estimated clusters, the 2D projection of the clusters using t-SNE approach, and the two activation maps which were related to the largest clusters are shown in this figure. As it can be seen in Fig. 5.6d, there are two major cluster of components, i.e. functional activation patterns, within the dataset. Convex hulls are generated around the estimated clusters. The components can be sorted based on the size of the clusters. As expected, the largest cluster (component) is associated with the activations in superior and middle temporal gyri that analyzes the low-level acoustic signals (Fig. 5.6e). The second largest cluster shows activations in the left inferior frontal gyrus which reflects higher-level linguistic (possibly semantic) processing of speech in the brain (Fig. 5.6f) [137, 142, 55, 161].

5.5 Discussions

In this study, we demonstrate that we can assess the reliability of the obtained components by a sparse representation analysis. Although here we used KSVD as a sparse representation analysis for an fMRI dataset, we an-
5.5. Discussions

Figure 5.6: Reliability and visualization analysis of the simulated data for two different sets of parameters (first and second row). (a) Number of times each component appeared in a cluster. (b) Dendrogram of the hierarchical clustering algorithm. (c) Similarities between estimated clusters, arranged according to the dendrogram. (d) 2D projection of the clusters. (e, f) Activation maps corresponding to the two largest clusters.

ticipate that this method will be useful for analyzing other neuroimaging datasets (such as EEG and MEG), and other sparse representation techniques (such as [139, 67, 100]) can be tried. Future work will apply this method in order to investigate the usefulness of the proposed approach for other datasets. Future work should also provide a quantitative measure for interpretation of clusters and tuning of the sparse representation parameters.

In summary, a method is proposed to investigate the reliability of the estimated components using sparse representation analysis. To achieve this goal the KSVD algorithm is run several times and estimated components are clustered based on their similarity. Then, the clusters are visualized using
5.5. Discussions

a nonlinear 2D projection. The proposed approach provides a tool for further investigation of the obtained sources. The approach highlights compact clusters with higher number of components, and suggests less reliable cluster of components that could be discarded.
Chapter 6

Multi-object Statistical Analysis of Major Depressive Disorder

6.1 Introduction

Depression directly affects more than 10% of the population at some point in their lives (World Health Organization, 2004), and is a leading cause of disability, with significant social, health and economic impacts [138]. Major Depressive Disorder (MDD) has a typical onset in adolescence and young adulthood, and prevalence rates of MDD by late adolescence equal those in adulthood [96]. MDD that starts in adolescence is associated with a large number of negative outcomes, including lower educational and occupational attainment, poor physical health, and poor interpersonal functioning [93]. These outcomes persist into adulthood and predict significant risk for a lifelong pattern of illness [18]. Given the enormous personal and societal costs associated with MDD, studies aimed at uncovering the pathology of the disorder in its earliest stages are crucial to informing effective prevention and intervention efforts.

Our understanding of the changes in brain neuroanatomy that are associated with MDD have benefited greatly from important advances in Magnetic Resonance Imaging (MRI) technology in the past two decades. Using structural MRI techniques in adult samples, differences in volume and shape have been found between depressed and non-depressed groups in temporal (e.g., superior temporal gyrus (STG)), hippocampus, amygdala), frontal (e.g., anterior cingulate cortex (ACC)), and orbitofrontal regions (see [110] for a review of the structural MRI findings associated with MDD in adulthood).

\[\text{This chapter is adapted from [153]: Mahdi Ramezani, Ingrid Johnsrude, Abtin Rassoulizadeh, Rachael Bosma, Ryan Tong, Tom Hollenstein, Kate Harkness, and Purang Abolmaesumi, Temporal-lobe morphology differs between healthy adolescents and those with early-onset of depression, Neuroimage: Clinical, 2014.}\]
These studies, conducted in adults, are likely to reflect the pathophysiology of MDD, as well as secondary changes due to longstanding behavioural alteration, and iatrogenic changes (as a result of pharmacological and other therapies).

To date, a small handful of studies have also investigated pediatric and adolescent-onset MDD and have reported structural differences from healthy controls in similar regions, including hippocampus [118], amygdala [162], striatum and caudate nucleus [122,173], superior and middle temporal gyri [173], and subgenual prefrontal cortex [22]. A compelling recent study by [37] even found volumetric differences in left hippocampus in clinically non-depressed young girls at high risk for depression (due to a maternal depression history), in comparison with girls who did not have a maternal depression history. However, other studies of early-onset depression have failed to find volumetric differences between depressed and healthy control groups in critical brain regions, including prefrontal cortex (e.g., [135], hippocampus [162], and amygdala [118]).

One potential reason for the failure to find consistent evidence of morphological differences in critical cortico-limbic circuits in early-onset MDD may be that such differences are subtle. Since the extent of hippocampal volume loss has been found to correlate significantly with the number of depressive episodes (i.e., time spent depressed) in adults with depression [128, 177], differences between depressed and non-depressed groups are likely to be larger in older samples of adults with recurrent depression than in younger individuals in the earliest stages of the illness. Hippocampal volume loss has also been associated with traumatic life events, which can be expected to accumulate with age (e.g., [38, 200]). More sensitive methods than have been used to date may be required to detect subtle differences in brain morphology associated with depression in its earliest stages, and in its youngest sufferers.

Previous methods used for investigating the morphological differences between individuals with depression and healthy controls can be categorized into three main types: 1) volume analysis; 2) analysis of local composition of tissue; and 3) analysis of shape and volume. The most common approach is hippocampal volume analysis using manual or automated segmentation [13, 196] [16]. In such analyses the volume of the hippocampal region is measured after isolating it from the rest of the brain. Using this method, several groups have observed smaller hippocampal volumes in adults with MDD [24, 26, 77, 120, 134, 169] whereas other groups have reported no differences or even larger hippocampal volumes [87, 131, 165, 200].

Voxel-based morphometry (VBM) [6, 81] which examines voxelwise dif-
ferences in grey- and white-matter volume and concentration throughout the brain, has demonstrated reduced grey matter intensity in hippocampus of MDD subjects [196, 16, 175, 36]. A limitation of VBM is that each individual’s brain data is normalized using nonlinear deformation fields to a reference template. Through that process, crucial idiographic information such as the shape of brain structures and their position, orientation and size (pose), both relative to other structures and in absolute terms, is lost [6]. This information may be critical for capturing group differences, particularly when such differences are subtle.

Alternatives to VBM approaches include Deformation Based Morphometry (DBM) [19] and Tensor Based Morphometry (TBM) [41], which are widely used to study the brains of people with schizophrenia, autism, dyslexia and Turner’s syndrome [72]. Unlike VBM, which analyzes images after the deformation fields have been applied in order to map any individual brain into a standard reference, these approaches take the deformation fields themselves as the dependent variable. Neither of these approaches has yet been attempted to study structural changes in depression. However, shape-analysis methods that are related to DBM/TBM have been employed in two separate studies to examine hippocampal differences in depression. These studies have focused on separate analysis of both shape and volume of the hippocampus using high-dimensional mapping [147] or spherical harmonic basis functions [204]. These studies with adult and elderly depressed participants reveal significant differences in hippocampal shape, but no volumetric differences. In these analyses, contribution to morphology made by the shape and pose of the hippocampal region and the surrounding regions was ignored.

Multi-object analysis enables the simultaneous statistical analysis of multiple brain structures, possibly allowing for the identification of subtle morphological differences across multiple brain regions, between groups. Multi-object methods were originally designed to characterize the shape of a population of geometric entities [63, 193, 111, 34], and have since been applied to analysis of brain MRI images to discriminate between healthy and clinical populations (e.g., pediatric autism; [82]), but has not yet been employed in the context of major depressive disorder.

In this chapter, we report the first use of a multi-object statistical pose and shape model to simultaneously analyze several temporal-lobe structures that have been implicated in MDD. Given that MDD is associated with morphological changes in several brain structures, pose and shape analysis of these brain structures simultaneously may be more sensitive to subtle group differences than is independent analysis of those structures, since simultaneous analysis includes information not just about the pose of brain
structures, but about their pose relative to each other. In this chapter, we first present the method, and then use it to identify temporal-lobe structures of interest and to characterize the relationship between the pose and shape of these structures and the symptomatology of early-onset MDD, when morphological differences between healthy and clinical groups are expected to be mild, and subtle. Use of a young sample at the earliest stage of their depressive illness has important implications for understanding the neurostructural correlates of the etiology of MDD.

6.2 Method

Pose and shape analysis of multiple brain structures, shown schematically in Fig. 6.1 involves three steps: a) preprocessing the MRI data to extract surface points on brain structures of interest; b) finding the pose and shape variations among these brain structures; c) Principal Component Analysis (PCA) on pose and shape variations in the subject population.

6.2.1 Preprocessing

The structural MRI data of the subjects are preprocessed using Statistical Parametric Mapping software (SPM8, Wellcome Department of Cognitive Neurology, London, UK). Briefly, Grey Matter (GM), White Matter (WM) and Cerebrospinal Fluid (CSF) are segmented using the automated segmentation processes in SPM. This results in a set of three maps for GM, WM and CSF in native space for each subject, in which each voxel is assigned a probability of being one of the three tissue types. The LONI Probabilistic Brain Atlas (LPBA40/SPM5) [176] in MNI space was used to extract left and right hippocampus, parahippocampal gyrus, putamen, and superior, inferior and middle temporal gyri from the brain of each participant (see Fig. 6.2; these are structures that have been shown to be associated with MDD in adulthood [110]). The LONI atlas is constructed using MRI data of 40 healthy volunteers, and 56 structures were labeled manually. We use the maximum-probability values at each voxel to segment the regions of interest in the atlas. To accomplish segmentations in each of the participants, we use the DARTEL algorithm to register the LONI atlas to each participant’s structural MRI, and extract surface points, $V = \{v_{n,l}\}_{n=1, \ldots , N, l=1, \ldots , L}$, indexing the coordinates of the surface voxels on each of the selected brain structures [5]. Here, $v_{n,l}$ consists of all surface points of the $l$th structure of subject $n$, $L = 12$ is the number of structures, and $N = 45$ is the number of subjects in the training set. For each subject, the surface boundary of each
6.2. Method

Brain structure was used to compute the volume of that structure. Structure volumes were compared between the MDD and control groups.

Figure 6.1: Schematic of the pose and shape statistical analysis of multiple brain structures. (a) Preprocessing the MRI data for extracting surface points on brain structures of interest; (b) Pose and shape multi-object analysis for finding the pose and shape variations between multiple brain structures; (c) PCA for generating pose and shape features.
6.2. Method

Figure 6.2: Segmented structures in both hemisphere of the brain which are used for multi-object statistical analysis. Surface points of putamen (blue), hippocampus (green), parahippocampal gyrus (red), ITG (cyan), MTG (yellow), and STG (magenta) in both hemisphere of brain are shown in (a) anterior to posterior view, and (b) posterior to anterior view. Structures in left hemisphere of the brain are shown in (c).

6.2.2 Pose and Shape Analysis

Since all surface points are extracted using the atlas in MNI space, the correspondences among the surface points (between homologous points in different subjects) was known. We used those correspondences to compute the linear (rigid plus scaling) deformation required to warp each structure in each participant to the mean shape of each structure calculated across participants, using generalized Procrustes analysis [62]. Pose variations were calculated using translation, rotation, and scaling values of these deformation fields. Each transformation for a voxel, \( x \), is defined as \( T(x) = sRv + d \), where \( R \) is a rotation matrix, \( d \) is a translation vector, and \( s \) is a scale factor. These transformations form a Lie group, which is a Riemannian manifold so conventional statistical analysis in Euclidean space is not applicable. However, a logarithmic transform was used to put the members of the Lie group into linear tangent space, appropriate for conventional statistical analysis. Exponential and logarithm maps are performed using the standard matrix exponential and matrix logarithm; e.g., the matrix exponential is defined by the series:

\[
exp(T) = \sum_{0}^{\infty} \frac{1}{k!} T^k
\]
6.2. Method

The logarithm of the transformation is defined as:

$$\log(T) = \begin{bmatrix} l & -r_z & r_y & x \\ r_z & l & -r_x & y \\ -r_y & r_x & l & z \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

(6.2)

where $l = \log(s)$, and $(r_x, r_y, r_z)$ is the rotation axis with angle $\theta = \sqrt{r_x^2 + r_y^2 + r_z^2}$.

Thus, each transformation, $T_{n,l}$, which represents the transformation from the $l$th structure in the mean shape to the corresponding structure in the $n$th instance, was expressed as a vector with seven variables: $(r_x, r_y, r_z, x, y, z, l)^T$.

For the purpose of statistical analysis, each transformation was normalized using the mean transformation for each structure, $M_l$, and mapped to the tangent space: $u_{n,l} = \log(M_l^{-1}T_{n,l})$ [20][143]. The transformation vectors were concatenated for each individual to form a $7L \times 1$ vector: $u_n = [u_{n,1}^T, \ldots, u_{n,L}^T]^T$ and the matrix of all transformations for all individuals was created: $U^p = [u_{1}^p, \ldots, u_{N}^p]^T$.

Shape variations are computed as the residual deformation required to map the mean shape of each structure to the corresponding structure for each subject, after the linear transformation for pose has been applied. Subsequently, similar to the pose variation extraction method described earlier, the distance vectors (deformations) were concatenated for each subject: $u_{n}^s = [u_{n,1}^s, \ldots, u_{n,L}^s]^T$ and the matrix of all transformations for subjects was created: $U^s = [u_{1}^s, \ldots, u_{N}^s]^T$.

6.2.3 Statistical Analysis

A multi-object statistical pose and shape model [21] was generated for the selected brain structures. In order to extract major directions of the pose and shape variations across all subjects, we constructed an orthonormal basis set that represented all pose and shape variations using Principal Component Analysis (PCA).

PCA on pose was performed using $U^p = A^p F^p T$. In this equation, $F^p = [f_1^p, \ldots, f_{N-1}^p]^{7L \times (N-1)}$ is the pose feature matrix, and $f_i^p$ are principal components that are sorted in descending order of their variance. $A^p = [a_1^p, \ldots, a_{N-1}^p]^{N \times (N-1)}$ is the corresponding weight matrix, generated from the principal component weights. We focus our analysis on the principal components associated with the pose that capture two standard deviations.
6.2. Method

of variations in the data. Similarly, PCA was used to identify an orthogonal vector set for shape, \( F^s = [f_1^s, \ldots, f_{N-1}^s] \), and the corresponding weight matrix, \( A^s = [a_1^s, \ldots, a_{N-1}^s] \). We consider principal components that capture one standard deviation of shape variations in the data. The primary difference between the number of principal components we consider for pose and shape stems from the difference in the dimensions of pose components (i.e. \( 7L \)) and shape components (i.e. the number of all surface points in each structure and is significantly larger than \( L \)).

Our objectives were to 1) identify pose and shape features that would differentiate the two groups; and 2) investigate the relation between these features and the clinical index of depression (i.e., BDI scores).

To achieve the first objective, we first use a random-forest classification [23] approach to sort the selected principal components. Random forests are a learning method for classification that use multiple decision trees for training. The decision tree splits the weights related to the considered principal components to maximize diversity among the subjects [46]. As a result, a tree with nodes and leaves is constructed, where its top node shows the weights with maximum separability. We perform unpaired two-sample t-tests (assuming unequal variance in the two groups) only on the top-node weights for pose and shape, i.e. one component for pose and one component for shape. As this study was designed to be hypothesis-generating and sensitive to morphological differences in brain structures between adolescent depressed individuals and control participants, a significance level of \( p < 0.05 \), uncorrected for multiple comparisons, was used [163, 89]. In order to visualize the significant pose component, associated with the top node weights, the norms of the three pose parameters (three translation, three rotation, and one scale variables) were computed. Subsequently, the mean of each parameter was removed and the result was divided by the standard deviation of the parameter. For the shape, the mean of the significant shape principal component associated with the top node weights was removed and the result was normalized to the component’s standard deviation. The higher the absolute value of the normalized pose or shape component is, the more the contribution of that member of the principal component is to capture the differences between the two groups.

To achieve the second objective, we calculate Spearman correlation coefficients between Beck Depression Inventory score and the top-node pose and shape weights.
6.3 Results

6.3.1 Volume Analysis

We first assessed the volume differences between the MDD and control groups for each structure. Unpaired two-sample t-tests (assuming unequal variance in the two groups) with significance level of $p < 0.05$, uncorrected for multiple comparisons, was used to detect volume differences between the two groups. Fig. 6.3 shows the distribution of the volume of each structure for the depressed and control groups.
Figure 6.3: Distribution of the volume of each structure between the two groups, (a) left putamen, (b) right putamen, (c) left hippocampus, (d) right hippocampus, (e) left parahippocampal gyrus, (f) right parahippocampal gyrus, (g) left inferior temporal gyrus, (h) right inferior temporal gyrus, (i) left middle temporal gyrus, (j) right middle temporal gyrus, (k) left superior temporal gyrus, (l) right superior temporal gyrus. The central red mark is the median, the edges of the blue box are the 25th and 75th percentiles, and the whiskers show the extreme values of the volumes.
The volume of both the left parahippocampal gyrus and the left superior temporal gyrus were significantly greater (p = 0.019 and p = 0.034 respectively) in the depressed than the control group.

### 6.3.2 Pose and Shape Analysis

The goal of our multi-object analysis was to investigate the pose and shape differences in brain structures between the participants with MDD and healthy controls. The first four principal components of pose capture two standard deviations (95\%) of the variation in pose, and the first eight components of shape capture one standard deviation (68\%) of the variation in shape. The random-forest classification trees for pose and shape were built on these components. Statistical analyses using unpaired two-sample t-tests were performed on the top component for each tree. The two groups differed significantly (p = 0.031 with corresponding statistical power of 0.77 [66] for the pose component, and p = 0.042 with corresponding statistical power of 0.89 for the shape component). Table 6.1 shows the normalized pose parameters across different brain structures for the most significant pose component. The translation component differed significantly between the two groups in left putamen, left and right hippocampus, and left ITG. Rotation also differed between the groups in left putamen, right hippocampus, and left and right ITG, and scale differed between groups in the left and right putamen.

Table 6.1: Normalized pose parameters of brain structures. L and R show the assigned anatomical left and right hemispheres.

<table>
<thead>
<tr>
<th>Structure</th>
<th>Left or right hemisphere</th>
<th>Translation</th>
<th>Rotation</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Putamen</td>
<td>L</td>
<td>1.16</td>
<td>1.93</td>
<td>1.81</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>0.66</td>
<td>0.28</td>
<td>2.32</td>
</tr>
<tr>
<td>Hippocampus</td>
<td>L</td>
<td>1.13</td>
<td>-0.07</td>
<td>-0.14</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>1.50</td>
<td>-1.11</td>
<td>0.03</td>
</tr>
<tr>
<td>Parahippocampal gyrus</td>
<td>L</td>
<td>-0.31</td>
<td>-0.75</td>
<td>-0.41</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>0.54</td>
<td>-0.12</td>
<td>-0.79</td>
</tr>
<tr>
<td>Inferior temporal gyrus</td>
<td>L</td>
<td>-1.82</td>
<td>-1.58</td>
<td>-0.52</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>-0.90</td>
<td>1.3</td>
<td>-0.4</td>
</tr>
<tr>
<td>Middle temporal gyrus</td>
<td>L</td>
<td>-0.73</td>
<td>-0.72</td>
<td>-0.28</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>-0.70</td>
<td>0.81</td>
<td>-0.55</td>
</tr>
<tr>
<td>Superior temporal gyrus</td>
<td>L</td>
<td>-0.14</td>
<td>0.01</td>
<td>-0.27</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>-0.36</td>
<td>0.03</td>
<td>-0.80</td>
</tr>
</tbody>
</table>
6.3. Results

Figure 6.4: Shape principal component that was significantly different between the two groups. The component is normalized by removing the mean and divided by its standard deviation. Inferior and superior view of the (a, b) left putamen, (c, d) right putamen, (e, f) left hippocampus, (g, h) right hippocampus, (i, j) left parahippocampal gyrus, (k, l) right parahippocampal gyrus. The color smoothly varies from black through red, orange, yellow and white, to show the minimum through maximum difference values. Left side of the pictures shows the left side of the brain, right shows right side, top is the anterior and bottom is the posterior.
6.3. Results

Figure 5. Shape principal component that was significantly different between the two groups. The component is normalized by removing the mean and divided by its standard deviation. Inferior and superior view of the (m, n) left superior temporal gyrus, (o, p) right superior temporal gyrus, (q, r) left middle temporal gyrus, (s, t) right middle temporal gyrus, (u, v) left inferior temporal gyrus, (w, x) right inferior temporal gyrus. The color smoothly varies from black through red, orange, yellow and white, to show the minimum through maximum difference values. Left side of the pictures shows the left side of the brain, right shows right side, top is the anterior and bottom is the posterior.

Figure 6.5: Shape principal component that was significantly different between the two groups. The component is normalized by removing the mean and divided by its standard deviation. Inferior and superior view of the (m, n) left superior temporal gyrus, (o, p) right superior temporal gyrus, (q, r) left middle temporal gyrus, (s, t) right middle temporal gyrus, (u, v) left inferior temporal gyrus, (w, x) right inferior temporal gyrus. The color smoothly varies from black through red, orange, yellow and white, to show the minimum through maximum difference values. Left side of the pictures shows the left side of the brain, right shows right side, top is the anterior and bottom is the posterior.
6.3. Results

Figs. 6.4, 6.5 shows the normalized shape component across different structures in the brain. As can be seen in the figure, many regions of all the examined structures show variations of the shape that are more than 1.96 (two standard deviations away from the mean), in both hemispheres.

To investigate the relation between the pose and shape weights that were significantly different between the two groups and BDI scores, Spearman correlation coefficients were calculated between the pose and shape values and BDI. The significant pose component correlated significantly with BDI (Spearman correlation: 0.38, p-value = 0.0086, slope: -0.039, intercept: 0.39), but the significant shape component did not (Spearman correlation: 0.15, p-value = 0.298, slope: -0.89, intercept: 8.8). Fig. 6.6 shows the distributions of the pose scores (6.6a) and shape scores (6.6b) across BDI. The four male subjects are identified with a circle.

Figure 6.6: Pose (a) and shape (b) scores that generated the significant difference between the MDD subjects and controls across the Beck Depression Inventory Index (BDI). Pose scores are significantly correlated to the BDI (Spearman correlation: 0.38, p-value = 0.0086, slope: -0.039, intercept: 0.39). Shape scores are not significantly correlated to the BDI (Spearman correlation: 0.15, p-value = 0.298, slope: -0.89, intercept: 8.8). A circle has been drawn around the data of male subjects.
6.4 Discussion

We conducted a statistical analysis of pose and shape information from several brain regions in order to examine whether the brains of individuals with early-onset MDD differ from those of healthy controls. Indeed, despite a rather small number of participants, we were able to observe statistically reliable differences in the medial temporal lobe regions, and we also determined that some features captured by the pose and shape analysis correlated with depressive symptomatology as measured by the Beck Depression Inventory. The sensitivity of this method may be related to its ability to capture differences in the spatial relationships among structures, not simply differences within an individual structure.

We observed volume differences in the left parahippocampal gyrus and the left superior temporal gyrus (STG) structures between the depressed group and the control group. The STG volume and GM density differences between the MDD and control subjects was previously shown by [200] and [174]. The individuals studied by these authors had been diagnosed with MDD at least two years earlier; so a later stage of the illness than the clinical group in the current study. Our results indicate that differences in STG are present right from the earliest stages of the disease.

The most significant component of the pose, highlighted in Table 6.1, showed that the left and right putamen, the left and right hippocampus, and the left and right inferior temporal gyri were more affected by MDD. The scale parameter of the right putamen is the only parameter that showed at least two standard deviations of variation. The translation mostly affected the left inferior temporal gyrus, and the rotation mostly affected the left putamen.

Shape analysis revealed that all examined structures, including putamen, hippocampus, parahippocampal gyrus, and superior, middle and inferior temporal gyri, differed between the two groups, suggesting that multiobject shape analysis is a sensitive tool for the examination of morphological differences in clinical samples. Moreover, within the most significant component of the shape, we identified regions that were at least two standard deviations away from the mean of that component, highlighting regions that were more affected by MDD.

Importantly, depressive symptomatology, as indexed by BDI scores, correlated with the pose of the structures (Fig. 6(a)). While the volume increase in the fusiform gyrus, cuneus and precuneus, have been previously shown to have association with BDI increase in MDD [101], we are the first to show that pose variations of multiple structures are also affected by MDD, and
correlate significantly with BDI.

The significant brain structural abnormalities seen here in early-onset depression are consistent with those observed in previous work \[118, 119, 117\]. However, MacMaster et al. only investigated volumetric differences between brain structures, after isolating each structure from the rest of the brain. Here, we have investigated the morphometric differences using simultaneous pose and shape analysis of multiple structures. As a result, we can capture differences due to the relationship among structures, and also differentiate between pose and shape morphometric differences.

The neural mechanisms underlying the observed morphometric differences in MDD have received empirical attention. Depression is associated with chronic dysregulation of the hypothalamic-pituitary-adrenal (HPA) axis with resulting chronic release of cortisol and other neurotoxic stress hormones [25]. Glucocorticoid neurotoxicity has preferential effects on hippocampal neurogenesis (e.g., [199, 167]). Indeed, in both preclinical and human clinical studies chronic stress and depression are associated with long-term changes in the hippocampus in the expression of genes involved in synaptic plasticity, such as brain-derived neurotrophic factor (BDNF; e.g., [102, 130]). Our results extend the state of the literature by suggesting that through the use of sensitive pose and shape analyses, the structural differences in MDD can be observed at the very initial stages of the illness, suggesting that they do not just emerge over the recurrent and chronic pathology of the disorder.

A concern about the method is the possible dependence on the quality of the segmentation. In this work, the segmentation comes from an atlas and the registration of atlas to the brains of the individual participants. A potential alternative is to manually segment the structures in individual brains prior to a group-wise registration. In future studies, we can also use polyaffine transformations in a logarithmic domain [4, 45], instead of similarity transformations for registration of multiple structures. An affine transform would further encompass anisotropic scaling and shearing.

Another concern is that we did not make any formal adjustments to correct for multiple comparisons, which potentially introduced a risk of false-positive results. Therefore, the p-values should be interpreted with caution [60]. The use of multiple comparisons corrections is often debated, because these corrections increase the chance of making type II errors that minimize truly important findings and require the use of large samples (which are often prohibitively expensive in neuroimaging research) to detect modest effect sizes [140]. As such, future studies with larger samples are needed to further validate these results.
6.4. Discussion

The current study investigated morphological variation in the pose and the shape of hippocampus and surrounding structures in early-onset MDD compared to control participants. Although a large number of previous studies have shown differences between MDD subjects and controls [196, 16, 175, 36, 24, 26, 77, 134, 169], ours is the first, to our knowledge, to simultaneously analyze multiple structures, and to separate pose and shape in morphological analysis. The value of the presented method is that it identifies structures of interest and characterizes types of differences (i.e. pose and shape) that can then be fed back into models/theories on etiology. In other words, what is more relevant than finding group differences is pinpointing the effect of underlying mechanisms that lead to MDD on brain structures and their interrelationships.

In summary, using multi-object statistical pose and shape analysis, we demonstrated brain morphological differences between adolescents and young adults with early-onset MDD and healthy control subjects. Relative pose and shape information of multiple structures in brain, which are usually disregarded, were shown to be important in capturing the group differences. Within this framework, the shape deformations were analyzed separately from rigid transformations and scale (i.e., the pose information). Therefore, we could identify the type of morphological differences (pose and shape). Within the simultaneous analysis of multiple structures the relative differences among structures were captured. The differences were more pronounced in the moderate and severely depressed participants. Moreover, morphological features (pose) significantly correlate with depressive symptoms across both normal and depressed participants.
Chapter 7

Fusion Analysis of Brain Shape Deformations and Local Composition of Tissue

7.1 Introduction

Studies of adults with primarily recurrent episodes of MDD have shown significant volumetric differences in temporal (e.g., Superior Temporal Gyrus [STG], hippocampus, amygdala) and frontal (e.g., Anterior Cingulate Cortex [ACC] and Orbitofrontal cortex [OFC]) brain regions relative to healthy controls (see [14, 110, 15] for reviews of the neuroanatomy and structural MRI findings associated with MDD). The most consistent finding in these studies is reduced hippocampal volume in adult patients with MDD compared to healthy controls. However, some studies have also failed to find group differences in hippocampal volumes [200, 87, 131, 165], and others have even reported larger hippocampal volumes in patients with MDD relative to healthy controls [77, 194, 200].

A large number of approaches have been developed to characterize differences, among individuals and groups, in the neuroanatomical configuration of the human brain. Generally, these approaches are classified into those that measure differences in brain shape, and those that measure differences in the local volume (and concentration) of brain tissue after macroscopic differences in shape have been discounted [6]. The former approaches analyze the deformation fields required to map individual brains onto some standard reference in order to characterize neuroanatomy. Deformation Based Morphometry (DBM) [19] and Tensor Based Morphometry (TBM) [41], are widely used approaches that use deformation fields. Shape-analysis meth-

\[^1\]This chapter is adapted from the following submission [151]: Mahdi Ramezania, Purgang Abolmaesumia, Amir Tahmasebih, Rachael Bosma, Ryan Tong, Tom Hollenstein, Kate Harkness, and Ingrid Johnsrude, Fusion Analysis of First Episode Depression: Where Brain Shape Deformations Meet Local Composition of Tissue, *Neuroimage: Clinical*, 2014.
7.1. Introduction

Methods that are related to DBM/TBM have been widely employed to examine morphometric differences in depression. For example, in MDD, \[147\] used high dimensional brain mapping on MRI data to quantitatively characterize the shape and volume of the hippocampus in adults with MDD and healthy controls (mean age = 33 ± 10). They found significant group differences in hippocampal shape, but no evidence for differences in volume. In a more recent study, \[204\] applied SPHERical HARMonic (SPHARM) shape analysis to the left and right hippocampi of elderly patients with MDD (age > 60) and healthy controls. Analysis revealed significant shape differences in the mid-body of the left hippocampus between the two groups. Further, in terms of volume, patients in a current episode of MDD had lower left hippocampal volumes in comparison to controls, whereas patients in remission from MDD showed no reduction in hippocampal volume. In previous chapter, we used multi-object statistical pose and shape analysis, and demonstrated brain morphological differences between adolescents with early-onset MDD and healthy controls.

Approaches that focus on the local composition of brain tissue, such as VBM, compare tissue images on a voxel-by-voxel basis after the deformation fields have been used to spatially normalize the images. For example, \[13\] applied VBM using SPM99, and reported smaller grey-matter volume of the right hippocampus, and smaller white-matter volume in the left anterior cingulate and right middle frontal gyrus, in elderly patients with MDD compared to healthy controls. Using VBM in SPM5, \[196\] reported significantly lower left hippocampal volumes in middle-aged patients with MDD in comparison to healthy controls. Similarly, in the same group of middle-aged patients with MDD, \[16\] compared VBM using a manual segmentation method and the automated method, and found significant hippocampal volume reductions using both segmentation methods in comparison with healthy controls. Finally, studies focusing on younger age groups, and including relevant covariates (i.e., age, sex, and intracranial volume) have also reported significantly lower hippocampal volumes, particularly in the left hemisphere, in both adolescents with MDD \[118\] and in patients with early onset MDD and a family history of depression \[117\].

In summary, computational neuroanatomical techniques either use the deformation fields themselves to characterize brain structural variation, or use these fields to normalize images that are then entered into an analysis of regionally specific differences in tissue composition. Ideally, a procedure like VBM should be able to automatically identify any structural abnormalities in a single brain image. However, even with many hundreds of subjects in a database, the method may not be powerful enough to detect subtle
abnormalities [6]. Recently, unified voxel- and tensor-based morphometry (UVTM) is proposed that uses locally adaptive combination of TBM and VBM to improve sensitivity [97]. UVTM is an extension of the Jacobian modulated VBM [53], which gives weights to VBM or TBM analysis based on registration confidence. In modulated VBM, voxel concentration is scaled based on the amount of deformation which was applied in the registration procedure. Although the motivation for multiplying the Jacobian determinant of transformations and the tissue segmentation probabilities is intuitive, it is not clear if the statistically significant regions resulting from VBM and TBM will match, although it is assumed to be. In addition, there has been no quantitative study on determining the optimal weight parameters based on the registration confidence. A more powerful procedure would be to use a voxel-wise multivariate approach. Within a multivariate framework, in addition to images of grey matter concentration, other image features such as white matter concentration, and the deformation fields calculated during the spatial normalization procedure can also be included [6]. Fusion of these multiple images may help in detecting subtle individual differences.

Joint Independent-Components Analysis (jICA) [30] is a multivariate technique for such "fusion analysis". It combines information from multiple features, which are a lower-dimensional representation of selected brain structures. jICA, as a group-level analysis technique, uses extracted features from individual subjects’ data and tries to maximize the independence among joint components. For example, [39] combined resting state functional connectivity and fractional anisotropy data within jICA in a dataset of four subjects with MDD and nine healthy control subjects to investigate links between functional connectivity changes and white-matter abnormalities. They reported differences in the strength of connectivity and in the coherence of white-matter tracts among subgenual anterior cingulate cortex (sACC) and perigenual ACC, anterior midcingulate cortex, caudate, thalamus, medial frontal cortex, amygdala, hippocampus, insula, and lateral temporal lobe.

The purpose of the current study was to combine, for the first time, brain shape and regional brain tissue composition using multivariate jICA technique in order to investigate the brain structural correlates of first-episode MDD. We determined the joint variation of shape and tissue composition in the hippocampal region in a sample of young people suffering from a first episode of MDD in comparison to a sample of young healthy controls. The importance of a young first-episode group is that they have not been subject to the known neurotoxic effects of glucocorticoids resulting from aging and the pathology of chronic depression [167,172]. We hypothesize that, whereas
conventional univariate analysis may not be sensitive to subtle differences in brain structure in this group, a multivariate technique that jointly analyzes multiple brain characteristics (i.e., shape and tissue composition) may have the requisite sensitivity to capture group differences. Following a group-wise registration using DARTEL [20, 16] to create an average template, we obtained individual grey matter (GM) and white matter (WM) tissue maps in the template space, along with the deformation fields required to warp the template to the GM and WM maps. Using the jICA technique, we combined these features, reflecting the tissue composition and shape of the brain in each individual, in order to extract spatially independent joint sources and their corresponding modulation profiles. We hypothesize that the mixing coefficients of the modulation profiles will lead to better discrimination of MDD subjects from the control group compared to the results obtained when brain shape and tissue composition are analysed separately.

7.2 Method

In the following two subsections, first the input features to the jICA method, representing tissue composition and deformation of selected brain structures, are described. Then, the multivariate joint independent-components analysis technique, used to fuse multiple features, is briefly reviewed.

7.2.1 Features

The data type on which we focus in this paper is structural MRI (sMRI). Outcome measures derived from structural images may include measures of shape (e.g., deformation) or tissue volume or concentration (e.g., grey or white matter). Below, we describe how we extracted three different features: (1) shape deformation information, and (2) grey- and (3) white-matter concentration used for voxel-based morphometric (VBM) analysis.

The sMRI data were preprocessed using Statistical Parametric Mapping software (SPM8, Wellcome Department of Cognitive Neurology, London, UK). Briefly, GM, WM, and cerebral spinal fluid (CSF) were segmented using the automated segmentation processes in SPM. This resulted in a set of three images in native space, in which each voxel is assigned a probability of being one of the three tissue types. The GM maps were registered using

---

2 The reason for using these inverse deformation fields from the template to each subject is that we can use the correspondences among the voxels (between homologous voxels in different subjects) to compute the jICA decomposition.
the DARTEL method, which achieves accurate inter-subject registration of images \cite{5, 16, 187}. The DARTEL procedure uses the GM and WM maps to create new templates and warps the GM and WM maps of each subject to the DARTEL template. Using DARTEL group-wise registration, the inter-subject registration is more accurate comparing to other SPM tools, therefore less spatial smoothing can be performed. We have used Gaussian convolution kernel with Full Width at Half Maximum (FWHM) of 8 mm. To demonstrate the effect of smoothing, we report the results with and without spatial smoothing. The deformation fields (DF) required for warping the groupwise (DARTEL) template to the GM and WM maps of each subject were also created. These deformation fields show how much a participant’s structure deviates from that of the other participants. The absolute value of the deformation field (displacement) for each voxel is used to represent shape morphometry. The warped GM and WM segments along with the deformation fields are input features to the joint analysis method.

To reduce the number of voxels in the analysis, a segmented LPBA40/SPM5 atlas \cite{176} in MNI space was used to extract the anatomical regions of interest. We selected the hippocampal region since abnormalities in this region have been associated with the pathology of MDD \cite{110, 14, 15}. To account for atlas-to-subject small registration errors, the selected region was dilated using a disk with the radius of 5 voxels, with morphological operators to include adjacent regions in addition to the selected brain structure. Voxels inside the created mask were selected for joint analysis.

7.2.2 Joint Independent Component Analysis

We assume that there is a relation between brain tissue type (GM or WM) differences and brain structural deformations. This is not an unreasonable premise: if depression is associated with differences in both the size and shape of brain structures, then differences in the volume and/or concentration of gray and/or white matter might be related to differences in structural deformations in depressed individuals relative to controls. The three features described in the previous section were used as input observations \((X = [x_1, x_2, \ldots, x_N]^T \in \mathbb{R}^{N \times K})\) to jICA in order to combine brain shape deformations and local composition of tissue. jICA can be used to identify any joint set of features \((S = [s_1, s_2, \ldots, s_N]^T \in \mathbb{R}^{K \times N})\) that is anatomically differentiable between depressed subjects and healthy controls, where \(x_i\) \((i = 1, 2, \ldots, N)\) is the vector of stacked features for subject \(i\), and \(s_i\) shows the \(i\)th joint independent component (source). \(N\) is the number of subjects and \(K\) is the total dimensionality of stacked vectors. Considering
7.2. Method

the generative model $\mathbf{X} = \mathbf{AS}$, the aim of jICA method is to find the matrix $\mathbf{W} = \mathbf{A}^{-1}$ so that the estimation of $\mathbf{U} = \mathbf{WX}$ is close to $\mathbf{S}$. In this model, $\mathbf{A}$ is the matrix of mixing coefficients (also called ICA loading parameters, or the modulation profile), and $\mathbf{W}$ is the unmixing matrix. A schematic of the jICA approach is shown in Fig. 3.1.

Joint independent components were found using the Infomax algorithm [12], which is based on minimization of mutual information of components. In this algorithm, the output entropy of a neural network is adaptively maximized with as many outputs as the number of Independent Components (ICs) to be estimated. In order to use ICA, it is necessary to first specify the number of Independent Components (ICs) expected. We first attempted to estimate the number of ICs using the Minimum Description Length (MDL) criterion, which is an information-theoretic technique for model order selection [105]. Using the MDL criterion, the number of components in GM and WM was estimated to be 4 and 3, respectively, but because of the heterogeneity in the location and extent of deformations across both groups, this information-theoretic criterion did not converge for the on deformation field dataset. Accordingly, we instead follow the precedent set by [180] and set the number of ICs equal to $\frac{1}{3}$ of the total number of subjects: so for 25 subjects here, we specify eight components.

Separability of the mixing coefficients was used as a criterion for capturing group differences. These low-dimension coefficients reflect how much each subjects shape deformation and tissue composition is modulated by a joint source. To investigate whether the mixing coefficients truly differ between groups, we used two-sample (unpaired) t-tests. We report mixing coefficients that differ significantly between the two groups ($p<0.05$), and for which the corresponding z-scaled component had more than 10 voxels with values above a threshold of $|z| > 2.5$ (99.4% cumulative probability). We followed the precedent set by [2] to select minimum number of voxels within a cluster, and [192, 185] to select the threshold.

In order to determine whether the fusion analysis is superior (in terms of sensitivity) to analyses based on single feature, we examined the mixing coefficients for subsequent follow-up analyses for 4, 10 and 12 components to further confirm the validity of our model and to test for the stability of the joint independent components. Stability analysis of the results for different number of independent components showed replication of findings for 10 and 12 independent components; however, using eight components yielded stronger group differences, and higher z-values. As expected [112], under-estimating the number of components (e.g., choosing four as the number of ICs in our case), yielded less reliable results. Results of analyses with 4, 10, and 12 ICs are available from the author by request.
coefficients and component maps of datasets containing single features. We compared the result of the t-tests on the mixing coefficients from the jICA of GM, WM and DF features to the result of the t-tests on the mixing coefficients from the ICA of each of the three features separately.

To further investigate the group differences, columns of the mixing coefficients matrix, which reflect the weighting of each joint source in a subject’s GM, WM and DF, were used as input features to a classification algorithm. A discriminant analysis with a quadratic discriminant function was used to classify the subjects. Performance of the classifier was measured using leave-one-subject-out cross-validation, averaging classification performance across iterations. The joint ICA classification result was compared to classification results obtained with one or two features. The mixing coefficients were used as input features for classification of depressed and control subjects.

Furthermore, separability of the joint source distributions was quantified by computing a divergence measure between joint histograms. Each of the joint sources was divided into three maps, which correspond to the GM, WM and deformation field features used in the jICA analysis. The map elements (each one representing a specific voxel) were thresholded and sorted in descending order by the voxel value, resulting in a set of voxels representing the greatest differences between groups in each joint source. For each subject, voxels that survived thresholding in all three maps were counted on a three-dimensional joint histogram in a bin defined by the three input feature values (from the input observation matrix $X$ in Fig. 3.1) at those voxels’ locations (see [31]) for more details on computing joint histograms). The group-averaged joint histograms were then calculated by taking the mean of the joint histograms across all the subjects in the group. The difference between the two groups was then assessed using the Renyi divergence formula [90]. The divergence was also computed for other combinations of features (two or one). The higher the values of the Renyi divergence criterion, the better the discrimination between groups [29]. The best combination of features is the one that yields the highest divergence value.

7.3 Experiments and Results

Structural brain differences are generally more apparent in patients with more severe or persistent forms of the illness [110]. Therefore, the analyses of this chapter were performed on a subset of 11 subjects (age: $18 \pm 0.89$, range: 16-21, 2 male, all right-handed) with moderate to severe levels of depression symptoms, as defined by a score of 19 or greater on the Beck Depression
7.3. Experiments and Results

Inventory (BDI-II). A similar-sized comparison group of 14 healthy controls (all 18 years old; all female, all right-handed) with BDI scores of zero were selected to act as the comparison group. The groups were well matched in age (p-value = 1). There was no socioeconomic status (SES) differences between the subjects in the two groups (p-value = 0.50).

We performed VBM analysis on GM and WM images obtained by DARTEL group-wise registration of the maps using SPM8 toolbox. We used the same explicit mask (described in 7.2.1) that we had used for the joint ICA analysis of multiple features. Results show no significant WM, or GM differences using Family Wise Error (FWE) rate of 0.05 or significance level of 0.001, and cluster size of more 10 voxels.

We report the statistical difference among joint sources to evaluate the performance of the proposed joint analysis. Two-sample t-test were performed on the mixing coefficients; i.e., the eight columns of matrix A, which correspond to eight independent components, where each column consists of two groups of coefficients (one for each group of participants). One source differed significantly between the two groups (p = 0.004, which passed the Bonferroni correction for multiple comparisons (p<0.00625)). Fig. 7.1a shows the mixing coefficients (i.e., weights) for this joint source, and its GM, WM and deformation-field components. The weights in the depressed group were significantly higher than in the control group. Figs. 7.1b, 7.1c, and 7.1d depict the statistical Z maps around the left and right hippocampus (the regions of interest) for this joint source, and Table 7.1 shows the corresponding stereotaxic coordinates in MNI space. As can be seen in Fig. 7.1d and Table 7.1, the shape variations appear mostly in the left hemisphere of the brain within the hippocampal region. On the hand, we observe that the changes in the GM and WM concentration appear in both hemispheres, as shown in Fig. 7.1b and 7.1c. We remind the reader that for each subject, within the jICA framework (see Fig 3.1), the coefficients that modulate the three maps (shape deformations GM and WM concentrations within each joint source) are the same. In other words, the three maps, which represent the variation of shape deformations, GM and WM concentrations among subjects, are jointly related. Hence, our results indicate that the statistically significant shape deformations observed within the left hemisphere of the brain in the hippocampal region are related to the statistically significant GM and WM alterations in the hippocampal region in both hemispheres. It is reasonable to infer from these results that local changes in brain tissue composition may lead to alterations of shape in distant regions, because brain is an interconnected organ.

To investigate the effect of using fusion analysis to determine the in-
dependent components, the result of joint analysis of GM+WM+DF, and separate analysis of each of the GM, WM, and DF were compared. Eight independent-sample t-tests were conducted to compare depressed and non-depressed groups on the columns of the mixing coefficients for joint or separate analysis of features. The modulation profiles differed significantly between the two groups (Table 7.2). Separate analyses of GM, WM and DF failed to identify significant group differences. Results confirm that combination of shape deformations and local composition of tissue, but neither shape nor local composition of tissue alone, can discriminate between individuals in the two groups. As it can be seen smoothing hasn’t affected the results by much.

Table 7.3 shows the average classification error for jICA (first column), and ICA (last three columns) of GM, WM, and DF, each used as input features in data fusion analysis. Results show that the control and depressed subjects can be classified based on structural MRI data with an error of 32% using the combination of shape deformations and tissue composition (GM+WM+DF). The classification error using shape deformations or tissue composition alone was more than 36%. Considering that the number of subjects is low and the dimensionality of the input MRI dataset is quite high, the results are very promising.

Fig. 7.2 shows the group-average marginal histograms for GM, WM, and deformation field, respectively. As it can be seen, the histograms of the normalized intensity values GM and WM were almost the same for the two groups, whereas the histogram of the absolute deformations showed around 0.3 mm more deformation for subjects with MDD comparing to healthy controls.
7.3. Experiments and Results

Figure 7.1: Joint Independent Component Analysis (JICA) of brain tissue composition and shape deformation. Fig. 7.1a shows the mixing coefficients for the depressed and control subjects wherein the central red mark is the median, the edges of the blue box are the 25th and 75th percentiles, and the whiskers show the extreme values of the coefficients. Figs. 7.1b, 7.1c, and 7.1d show the joint source map of the most significant component for (b) GM, (c) WM, and (d) deformation field. The green dots indicate the boundaries of the region of the interest which was created by dilating a mask around the hippocampus.
### 7.3. Experiments and Results

Table 7.1: MNI coordinates for the most discriminative source map in three contrasts. Voxels which are above a threshold of $|Z| > 2.5$, and create a cluster volume of more than 10 voxels, are shown in the table. L and R show the assigned anatomical left and right hemispheres, the coordinates and value of the maximum Z are also provided in the table. Not significant regions are shown by ns.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Volume (voxels)</th>
<th>random effects: Max Value (x, y, z)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L</td>
<td>R</td>
</tr>
<tr>
<td><strong>GM concentration</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>69</td>
<td>55</td>
<td>5.5(-33, -28, -12)</td>
</tr>
<tr>
<td>44</td>
<td>74</td>
<td>5.0(-39, -4, -26)</td>
</tr>
<tr>
<td>26</td>
<td>30</td>
<td>4.3(-30, 2, -27)</td>
</tr>
<tr>
<td>23</td>
<td>9</td>
<td>5.0(-41, -6, -26)</td>
</tr>
<tr>
<td>4</td>
<td>19</td>
<td>3.9(-29, -9, -18)</td>
</tr>
<tr>
<td>2</td>
<td>18</td>
<td>6.4(-39, -4, -27)</td>
</tr>
<tr>
<td>Negative</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>25</td>
<td>6.5(-36, 3, -29)</td>
</tr>
<tr>
<td>28</td>
<td>51</td>
<td>6.4(-39, -36, -14)</td>
</tr>
<tr>
<td>40</td>
<td>19</td>
<td>5.3(-36, -37, -12)</td>
</tr>
<tr>
<td>33</td>
<td>11</td>
<td>7.1(-38, 6, -26)</td>
</tr>
<tr>
<td>7</td>
<td>31</td>
<td>3.6(-26, 8, -20)</td>
</tr>
<tr>
<td><strong>WM concentration</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>77</td>
<td>27</td>
<td>7.8(-36, 3, -27)</td>
</tr>
<tr>
<td>42</td>
<td>73</td>
<td>8.0(-39, -36, -14)</td>
</tr>
<tr>
<td>16</td>
<td>72</td>
<td>3.4(-27, 2, -27)</td>
</tr>
<tr>
<td>55</td>
<td>31</td>
<td>6.5(-36, -37, -12)</td>
</tr>
<tr>
<td>28</td>
<td>18</td>
<td>8.7(-38, 6, -26)</td>
</tr>
<tr>
<td>Negative</td>
<td></td>
<td></td>
</tr>
<tr>
<td>61</td>
<td>91</td>
<td>6.1(-39, -4, -26)</td>
</tr>
<tr>
<td>80</td>
<td>86</td>
<td>6.6(-33, -28, -12)</td>
</tr>
<tr>
<td>30</td>
<td>15</td>
<td>6.2(-41, -6, -26)</td>
</tr>
<tr>
<td>18</td>
<td>1</td>
<td>7.8(-39, -4, -27)</td>
</tr>
<tr>
<td>2</td>
<td>16</td>
<td>4.5(-33, -33, -9)</td>
</tr>
<tr>
<td>13</td>
<td>11</td>
<td>5.2(-30, 2, -27)</td>
</tr>
<tr>
<td><strong>Deformation Fields</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>515</td>
<td>0</td>
<td>4.3(-45, -5, -18)</td>
</tr>
<tr>
<td>13</td>
<td>0</td>
<td>2.6(-32, -21, -24)</td>
</tr>
</tbody>
</table>
7.4 Discussion

Table 7.2: P-values of the most significant joint source, obtained from two-sample t-tests performed on the columns of the mixing coefficients generated by jICA (first three columns), and ICA (last three columns). First and second row show the results without and with spatial smoothing of the features. GM: gray matter; WM: white matter; DF: deformation field, each used as input features in data fusion analysis. The p-values displayed in the first three columns passed a Bonferroni correction for multiple comparison (p<0.00625).

<table>
<thead>
<tr>
<th>Combination</th>
<th>GM+WM+DF</th>
<th>GM+DF</th>
<th>WM+DF</th>
<th>DF</th>
<th>GM</th>
<th>WM</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>0.004</td>
<td>0.005</td>
<td>0.004</td>
<td>0.069</td>
<td>0.067</td>
<td>0.081</td>
</tr>
<tr>
<td>FWHM: 8 mm</td>
<td>0.005</td>
<td>0.006</td>
<td>0.005</td>
<td>0.069</td>
<td>0.087</td>
<td>0.036</td>
</tr>
</tbody>
</table>

Table 7.3: Classification error obtained from discriminant analysis of the mixing coefficients generated by jICA (first column), and ICA (last three columns). GM: gray matter; WM: white matter; DF: deformation field, each used as input features in data fusion analysis.

<table>
<thead>
<tr>
<th>Combination</th>
<th>GM+WM+DF</th>
<th>DF</th>
<th>GM</th>
<th>WM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error</td>
<td>32%</td>
<td>36%</td>
<td>36%</td>
<td>40%</td>
</tr>
</tbody>
</table>

The sorted maximum Renyi divergence values for different combinations of contrasts are shown in Fig. 7.3. Higher values indicate better discrimination between the groups. Therefore, as indicated in the figure, combining the deformation field and tissue composition yielded greater discrimination than utilizing either deformation field or tissue composition data alone. In particular, combining GM and DF yielded the highest level of discriminatory power in both samples.

7.4 Discussion

The current study is the first to report that joint analysis of brain shape and tissue composition is sensitive enough to identify subtle significant differences between young people in a first episode of MDD and healthy controls,
7.4. Discussion

Figure 7.2: Group-average histogram for the whole dataset on GM (a), WM (b), and Deformation field (c). The difference between histograms of the two groups in deformation field was more than GM and WM.

whereas separate analysis of shape and tissue composition fails to discriminate the groups. The identified corresponding sources demonstrate MDD-related links between WM, GM and shape deformation changes in the hippocampus, which were not detectable with univariate voxel-based methods. Assuming that the features share the same mixing coefficient matrix (modulation profile), jICA uses more information to estimate the same number of mixing coefficients and can therefore improve source estimations compared to ICA. The observed shape deformations in left hippocampus are related to GM and WM alterations in hippocampus in both hemispheres (see Fig. 7.1 and Table 7.1). These significant shape deformation differences in the left hippocampus are consistent with a previous study of shape [204], and volume
7.4. Discussion

Figure 7.3: Renyi divergence criteria values for different combination of features on differentiation between histograms. The first six highest value combinations are shown in the figure. The higher the values of the Renyi divergence, the better the discrimination between groups.

differences [196] in late-life MDD; and volume differences in adolescents with MDD [118]. Our results provide compelling evidence that shape-deformation differences in the hippocampus between depressed and healthy individuals are present to at least some extent even in the very initial stages of the illness; they do not simply emerge over the recurrent and chronic pathology of the disorder, and they are independent of any potential neurotoxic effects of chronic anti-depressant usage.

Results demonstrate that individuals can be classified relatively accurately (with 68% accuracy) into control and depressed groups by using only structural MRI data. This is consistent with previous studies on diagnostic classification of MDD using brain structural neuroanatomy (67.6% diagnostic accuracy reported by [47] and 77.8% prognosis accuracy reported by [136] using adult subjects). However, classification results reported using functional Magnetic Resonance Imaging are higher (94.3% reported by [203], 90.6% reported by [113], and 95% reported by [49]), suggesting that functional analysis of MDD is more suitable for diagnostic classification.

The group-average histograms for individual features indicated that among individual features, the deformation field was able to best discriminate between the two groups. However, the combination of GM and deformation
field captured the group differences better than any individual feature alone, or any other combinations of features, as indicated by the values of the Renyi divergence. These results suggest that future studies should use both deformation fields, used to normalize individuals’ brain to the reference space, and regionally specific analyses, such as tissue composition measures, to better understand the brain basis of MDD and capture structural differences between subjects with MDD and healthy controls.

The proposed method based on fusion of brain tissue composition and shape deformation successfully captured the differences in hippocampal shape and tissue composition between young people in a first episode of depression and healthy control subjects. Specifically, using the jICA method, significant shape deformation differences in the left hippocampus were observed between depressed and control groups. In contrast, no differences were detected between the two groups when a separate analysis of each feature was conducted. These results suggest that the jICA method may be a more sensitive technique for detecting morphological differences in brain tissue; such sensitivity may be particularly helpful when the sample size is relatively small, or when structural abnormalities are relatively subtle (such as in groups of young people who are very early in their disease course). The current results have important clinical implications. Although prospective studies with individuals at risk for MDD are needed to determine the causal role of these structural differences in MDD, the current results suggest that hippocampal volume loss may be a vulnerability factor for a particularly severe manifestation of MDD in the first onset.
Chapter 8

Simultaneous Analysis of Pose, Shape, and Tissue Composition

8.1 Introduction

Major depression has been linked to brain structural changes using five different types of approaches: 1) volume analysis of a single brain structure (region of interest) using manual or automated segmentation [13, 16, 24, 26, 77, 87, 120, 131, 134, 165, 169, 196, 200]; 2) separate analysis of shape and volume of a region of interest using high-dimensional mapping [147] or spherical harmonic basis functions [204]; 3) analysis of local composition of tissue in a single brain structure or whole brain using voxel-based morphometry (VBM) [16, 36, 175, 196]; 4) analysis of pose (i.e. position, orientation, and size) and shape of multiple brain structures [156, 157]; and 5) joint analysis of shape deformations and local composition of tissue [151].

Limitation of the region of interest based methods (first and second approaches mentioned above) is that analysis is performed only on a single brain structure after isolating it from the rest of the brain, and the relative shape and pose information between that region and the surrounding regions for the purpose of group analysis is largely ignored.

Limitation of VBM is that each individual’s brain data must be normalized to a reference template, and the proportion of Grey Matter (GM) concentration and absolute volume of corresponding voxels are compared across individuals. Through that process, crucial idiographic information from the spatial normalization procedure, such as shape, relative pose of brain structures with respect to each other, and their absolute pose, is lost [6]. This information may be critical for capturing group differences, particularly when such differences are likely to be subtle.

\footnote{This chapter is adapted from [159]: Mahdi Ramezani, Abtin Rasoulian, Tom Hollenstein, Kate Harkness, Ingrid Johnsrude, and Purang Abolmaesumi, Joint source based analysis of multiple brain structures in studying major depressive disorder, SPIE Medical Imaging: Image Processing, San Diego, US, 2014.}
As shown in chapter 6, pose and shape analysis of multiple brain structures, alleviates several issues associated with previous techniques, and allows for identifying subtle shape and volumetric differences across multiple brain regions between groups [156, 157]. However, simultaneous analysis of pose and shape variations within a multivariate framework has not been investigated yet.

In chapter 7, within the jICA technique, we showed that local changes in brain tissue composition (i.e., GM and WM concentrations) result in modulation of shape deformations in distant regions (in both control and MDD groups, although to different degrees). However, this approach has not been applied for analysis of multiple brain structures.

In this chapter, we use the joint Source-Based Analysis (jSBA) framework to identify common information across different brain structural MRI features, for classification of individuals with and without MDD. Here, a source includes regions of brain that together exhibit intersubject covariance and group differences. The framework consists of three components, as shown in Fig. 8.1: 1) feature generation, 2) joint group analysis and 3) classification of individuals based on the joint analysis results. In the proposed framework, information from pose, shape, and tissue composition of a selected brain structure are represented as features. For each individual, features are used within the joint group analysis to generate joint sources and their corresponding modulation profiles. Modulation profiles are used to classify individuals into different categories.

To the best of our knowledge, this is the first framework for quantitative classification of individuals with MDD based on simultaneous analysis of pose, shape, and tissue composition, obtained from multiple brain structures. Our key contribution is extracting multiple information from structural MRI data of each subject, and creating features of pose, shape and tissue composition of multiple brain structures for the joint group analysis. An added value of our method is in identifying structures of interest, and characterizing types of differences that can then be injected back into models/theories on etiology.

The proposed framework is evaluated on data from a group of subjects diagnosed with severe or moderate MDD. In a cross-validation leave-one-subject-out experiment, we demonstrate that the framework enables the classification of these subjects with around 70% accuracy, solely based on their structural brain data.
8.2 Method

The jSBA framework consists of three components, as shown in Figure 8.1: 1) feature generation, 2) joint group analysis, and 3) classification of individuals based on the joint analysis results.

8.2.1 Feature Generation

The primary outcome measure derived from a structural image may include a measure of a particular structure (e.g., deformation) or a description of tissue type (e.g., Grey Matter (GM) and White Matter (WM) concentrations). Below, we describe how we extract three different features: (1) local brain tissue compositions, (2) pose variations, and (3) shape deformations; represented by the superscripts “g”, “p”, and “s”, respectively.

The structural MRI data were preprocessed using Statistical Parametric Mapping software (SPM8, Wellcome Department of Cognitive Neurology, London, UK). Each voxel of each individual structural (T1-weighted) MRI was assigned a probability of being Gray Matter (GM), White Matter (WM) and Cerebral Spinal Fluid (CSF), using the automated segmentation processes in SPM. The GM and WM maps were used within the DARTEL groupwise registration method [5] to create a population template, and the deformation fields required to map the GM maps from each participant to the template space. The mapped GM segments were then spatially normalized to stereotaxic MNI space, where a segmented LPBA40/SPM5 atlas [176] was used to segment multiple structures in the brain, including hippocampus, parahippocampal gyrus, putamen, and superior, inferior and middle temporal gyri, from both hemispheres of the brain. For each individual, the normalized mapped GM concentrations for the voxels within the selected brain structures were used as tissue composition feature, and the matrix of all GM concentrations was created: \( Y_g = [y_{g1}^N, \ldots, y_{gN}^N]^T \).
8.2. Method

Using the DARTEL algorithm, the atlas was registered to each individual’s structural MRI, and deformation fields for such registrations were created. The deformation fields were used to warp the surface points of each of the selected brain structures to each individual’s brain volume. Pose features were calculated by the parameters of similarity transformations (translation, rotation, and scale) between surface points of the brain structures across subjects [158]. For each brain structure, the correspondences among the surface points across subjects was used to compute the similarity transformations \( T_{n,l} \) from the mean shape \( \mu_l \) of that structure across subjects to that structure in each subject. \( T_{n,l} \) represents the transformation from the \( l \)th anatomical structure in the mean shape to the corresponding one in the \( n \)th subject. These transformations form a Lie group, therefore a logarithmic mapping transforms them to a linear tangent space, where conventional statistical analysis can be applied. Each transformation was normalized using the mean transformation for each anatomical structure, \( M_l \), and mapped to the tangent space: \( y_{p,n,l}^p = \log(M_l^{-1}T_{n,l}) \) [143]. The transformation vectors were concatenated for each instance to form a vector: \( y_{n,l}^p = [y_{n,1,l}^p, \ldots, y_{n,L,l}^p]^T \) and the matrix of all transformations for subjects, representing pose variations, was created: \( Y_p = [y_{1,l}^p, \ldots, y_{N,l}^p]^T \).

Shape features are computed as the residual deformation required to map the mean shape of each structure to the corresponding structure for each subject, after the similarity transformation is applied. The distance between the mean shape of each structure across all individuals in the data set, \( \mu_l \), and surface points of each structure in each individual were computed. The transformation vectors were concatenated for each instance: \( y_{n,l}^s = [y_{n,1,l}^s, \ldots, y_{n,L,l}^s]^T \) and the matrix of all transformations for instances, representing shape deformations, was created: \( Y_s = [y_{1,l}^s, \ldots, y_{N,l}^s]^T \).
8.2. Method

8.2.2 Group Joint Analysis

The three features described in the previous section were used as input observations to the proposed joint sparse representation method \[150\], which is a dictionary learning algorithm based on K-SVD algorithm \[1\]. Figure 8.2 shows the schematic of this method. In this figure, \(Y^T = [Y_g, Y_p, Y_s] \in \mathbb{R}^{V \times N}\) is the observation matrix, where \(Y_g\), \(Y_p\), and \(Y_s\) are matrices of GM, pose, and shape features, \(X = [x_1, x_2, \ldots, x_N] \in \mathbb{R}^{K \times N}\) is the sparse modulation matrix, and \(D = [d_1, d_2, \ldots, d_N] \in \mathbb{R}^{V \times K}\) is the dictionary containing \(K\) signal atoms representing the joint maps of GM, pose and shape. \(V\), \(N\), and \(K\) are the number of variables (GM voxels + pose features + shape features), subjects and brain joint sources, respectively. Assuming that maps of joint sources share the same sparse coefficients matrix, the method tries to represent a large set of brain features as a sparse linear combination of a small set of joint sources (‘basis maps’). To assess the reliability of the estimated joint sources as well as for visual inspection of sources in low-dimensional space, we use the proposed reliability analysis and visualization method in chapter 5. To accomplish this, the parameters of the sparse representation analysis are set and the analysis is run multiple times, using the selected parameters. The estimated components are clustered based on their similarities using the hierarchical clustering algorithm. Finally the clusters are visualized using the t-SNE approach. The estimated joint sources and sparse coefficients corresponding to centrotypes of the clusters are used in the third component of the jSBA framework, i.e., classification.

8.2.3 Classification

In order to overcome the classification problems caused by high dimensionality of data and the small set of available subjects, columns of the sparse coefficients matrix, \(X\), which reflect the weighting of each joint source in a subject’s feature, were used as input features to a random-forest classifier \[23\]. Random forests are a learning method for classification that use multiple decision trees for training. The decision tree splits the weights related to the considered joint sources to maximize diversity among the subjects \[46\]. As a result, a tree with nodes and leaves is constructed, where its top node shows the weights with maximum separability.

Performance of the classification was measured by the leave-one-subject-out cross-validation and averaging classification performance on each left out data. In other words, the data were split into training subjects and a test subject. In the training phase, the jSBA was performed on training subjects.
This produced a new sparse coefficient matrix $X_{\text{train}}$ and joint source matrix $D_{\text{train}}$ that modeled the generation of the training observations $Y_{\text{train}}$. The rows of the sparse matrix $X_{\text{train}}$ were used as input features to train the classifier, which divided the training subjects into two classes of with and without MDD. The input features for the test subject, rows of $X_{\text{test}}$, were then found by OMP solution of $Y_{\text{test}} = X_{\text{test}}D_{\text{train}}$. The positions of these vectors in k-dimensional feature space, relative to the hyperplane found in running the classifier on the training set, determined the classification of the test subject.

## 8.3 Results and Discussion

The goal of our jSBA framework is to examine whether joint sources can be used to accurately distinguish two groups of subjects on the basis sMRI data. The success of this analysis is evaluated by examining the statistical difference among joint sources, and by a classification method. Fig. 8.3 shows the results of reliability and visualization analysis on the 45 subjects with and without depression. The number of times each source appeared in a cluster, the dendrogram of the hierarchical clustering algorithm, the similarities between estimated clusters, the 2D projection of the clusters using t-SNE approach are shown in this figure. As it can be seen, there are two major cluster of components, within the dataset. Convex hulls are generated around the estimated clusters. The components can be sorted based on the size of the clusters.

Two-sample t-tests on the mixing coefficients was performed, and two components were found to differ significantly ($p < 0.05$) between the two groups. These components showed different sparse coefficients ($p = 0.033$ and $p = 0.036$) in subjects with and without MDD. Figures 8.4a and 8.4b shows the distribution of the sparse coefficients for these two significant sources. These sparse coefficients reflect how much each subject’s structural MRI features (i.e. GM concentration, pose, and shape) are modulated by a joint source. Coefficients of the first significant different source (lowest p-value) were almost zero for the control subjects. This demonstrates that the features are modulated by the subjects with MDD more than controls.

Classification results show that individuals with and without MDD can be classified with an accuracy of 67% based on information gathered from multiple brain structures. Considering the fact that the number of subjects is low and the dimensionality of the dataset is high, the results are very promising.
8.3. Results and Discussion

Figure 8.3: Reliability and visualization analysis of the 45 subjects with and without depression. (a) Number of times each component appeared in a cluster. (b) Dendrogram of the hierarchical clustering algorithm. (c) Similarities between estimated clusters, arranged according to the dendrogram. (d) 2D projection of the clusters.

The current study is the first to report classification of subjects with MDD and healthy controls based on pose, shape, and tissue composition obtained from sMRI data. All participants in the MDD group were in their very first episode of depression and were medication-free, which may provide important clues as to the disorder’s initial etiology. Results confirmed that shape deformation differences in the selected brain structures between depressed and healthy individuals are present to at least some extent even in
the very initial stages of the MDD. In the current study, all of the depressed subjects were diagnosed to have severe and moderate depression, further studies on subjects with mild depression are required.

In summary, using the proposed jSBA framework, independent joint sources of GM concentrations, pose, and shape deformations were obtained and simultaneously analyzed, to capture the group differences. The framework was evaluated on brain MRI of young adults with, and without, MDD, and determined significant differences between these two groups. In a cross-validation leave-one-subject-out experiment, we demonstrate that the framework enables the classification of these subjects with 67% accuracy, solely based on their structural brain data.

Figure 8.4: Sparse coefficients for the significant joint sources ($p = 0.033$, $p = 0.036$) for MDD subjects and healthy controls (a, b), respectively; the central red mark is the median, the edges of the blue box are the 25th and 75th percentiles, and the whiskers show the extreme values of the coefficients.
Chapter 9

Conclusion and Future Work

9.1 Conclusion

In this thesis we proposed a framework that can distinguish patients and healthy controls when the number of available subjects is low, and the between group differences are subtle. To this end, in the third chapter we proposed to use multi-task fMRI data, and take advantage of the complementary information that exist among the tasks for group classification. We used the joint ICA method together with an SVM classification algorithm, and demonstrated that cognitive patterns can be used to classify individuals in the absence of behavioral differences. We showed that by combining multiple functional contrasts, revealing different patterns of brain activity, the overall performance of the classification improves.

In the fourth chapter, we proposed joint sparse representation analysis which is more appropriate for joint analysis of brain functional imaging datasets. We showed that the proposed analysis can effectively identify the common and unique information among different levels of brain cognitive patterns in multi-task fMRI data within different groups. To demonstrate the potential of the proposed framework, analyses of simulated fMRI data were performed followed by analyses of experimental fMRI data from normal subjects performing speech comprehension tasks. Simulations showed the superiority of the proposed method to the jICA for multi-task fMRI data analysis. Results on the experimental fMRI data also demonstrate that the jSRA method can better capture the brain functional activation patterns, and therefore the differences, between two groups.

In Chapter Five, we proposed a method to investigate the reliability of the estimated components using sparse representation analysis. To achieve this goal the KSVD algorithm was run several times and estimated components were clustered based on their similarity. Then, the clusters were visualized using a nonlinear 2D projection. The proposed approach provides a tool for further investigation of the obtained sources. The approach highlights compact clusters with higher number of components, and suggests less reliable cluster of components that could be discarded. Future work will in-
investigate the use of the modulation profiles, which are substantially compact and sparse, for reliable classification of individuals based on neuroimaging data.

In Chapter Six, we investigated the simultaneous analysis of multiple brain structures within the multi-object statistical pose and shape analysis. Relative pose and shape information of multiple structures in brain, which are usually disregarded, were shown to be important in capturing the group differences. Within this framework, the shape deformations were analyzed separately from rigid transformations and scale (i.e., the pose information). Therefore, we could identify the type of morphological differences (pose and shape). Within the simultaneous analysis of multiple structures the relative differences among structures were captured.

In Chapter Seven, we fused brain tissue composition and shape deformation which previously haven’t been integrated or fused. Using the jICA method, we successfully captured the differences in hippocampal shape and tissue composition between young people in a first episode of depression and healthy control subjects. Specifically, using the jICA method, significant shape deformation differences in the left hippocampus were observed between depressed and control groups. In contrast, no differences were detected between the two groups when a separate analysis of each feature was conducted. These results suggest that the jICA method may be a more sensitive technique for detecting morphological differences in brain tissue—such sensitivity may be particularly helpful when the sample size is relatively small, or when structural abnormalities are relatively subtle (such as in groups of young people who are very early in their disease course).

In Chapter Eight, we used the proposed jSBA framework, to jointly analyze GM concentrations, pose, and shape deformations to capture the group differences of young adults with, and without, MDD. In a cross-validation leave-one-subject-out experiment, we demonstrate that the framework enables the classification of these subjects with 67% accuracy, solely based on their structural brain data.

9.2 Future Work

The results reported above for sMRI analysis of subjects with and without depression should be interpreted in the context of the following limitations. First, although current results (Fig. 6.6) did not reveal any differences in the structures we examined between male and female individuals with MDD, the study comprised young women with moderate and severe depression al-
most exclusively; therefore, generalization to young men in a first episode of MDD, and young men and women with milder levels of depression severity requires further study. Future studies should also investigate whether the present results generalize to adults with recurrent MDD, as well as younger children with MDD.

Second, we did not assess for the presence of comorbid anxiety disorders or specific subtypes of MDD. Future studies are required to examine variation in brain morphology with differing depression syndromes in order to identify biomarkers of more homogeneous endophenotypes.

Third, it will be important in future research to determine whether the present results generalize to children and early adolescents with MDD, as significant corticolimbic plasticity remains throughout childhood and early adolescence [80], which may obscure any potential toxic effects of depression vulnerability.

Fourth, as all of our depressed participants were medication naive, the structural differences are not due to any potential neurotoxic effects of chronic anti-depressant usage. As such, the structural differences may emerge as a result of premorbid epigenetic vulnerabilities. For example, hippocampal volume differences have been shown in individuals with particular polymorphisms of genes known to impart risk for depression, but only in the context of environmental adversity, such as a history of childhood trauma (e.g., [188]), or maternal depression (e.g., [37]). Future prospective, longitudinal studies that follow children at risk of depression as a result of these vulnerabilities through to the onset of syndromal MDD are required to clarify the precise etiological and pathological mechanism underlying the relation of hippocampal volume loss to MDD.

Fifth, few recent studies in MDD have focused on the White Matter (WM) integrity using Diffusion Tensor Imaging (DTI), to assess the structural connectivity of the network between healthy controls and MDD subjects. Korgaonkar et al. (2014) showed structural connectomic alterations between nodes of the default mode network and the frontal-thalamo-caudate regions in 95 MDD outpatients comparing to 102 matched control subjects [99]. However, Choi et al. (2014) found no significant differences in WM integrity disruption between 134 medication-free MDD patients and 54 healthy controls [40]. Future studies may use multivariate approaches to include analysis of geometric changes (pose and shape), tissue concentrations (WM and GM) and structural connectome (DTI). Future studies should also consider differences between MDD and control groups in other brain structures of relevance to MDD.

Finally, the participants in this study were volunteers and, thus, may not
be entirely representative of the population of young people with depression. Nevertheless, as a community sample, they may be more representative than the subjects of most previous studies, which have relied on treatment-seeking patients in tertiary care centers.

In future, we will apply the jSBA framework for improving the clinical diagnosis of the brain disorders using multi-modal datasets. In collaboration with sexual health research laboratory of Queen’s University, we have gathered structural and functional MRI images of women with Provoked Vestibulodynia (PVD). We will analyze structural and functional imaging data from patients with PVD within the jSBA framework, by combining brain regional tissue composition, and functional patterns of activation. In addition, using multi-object statistical analyses of brain pose and shape deformations [153], we will investigate shape deformations and relative anatomical pose information that is usually discarded in the alignment process. The combination of brain structure and function data should make clearer the pathologic role of abnormalities, and reveal the joint information between the features, such as the GM/WM involvement associations with brain activity patterns. The analysis would likely improve our understanding of the relation between clinical scores, brain structural and functional changes.
Bibliography


Bibliography


Bibliography


Bibliography


[195] Gael Varoquaux, Merlin Keller, J-B Poline, Philippe Ciuciu, and Bertrand Thirion. Ica-based sparse features recovery from fmri


