Aix*Marseille université
Socialement engagée

NNT : 0000AIXM0001

THÈSE DE DOCTORAT

Soutenue à Aix-Marseille Université
le 25 Juin 2024 par

PHAM Minh Tuan

Les contributions des réseaux neuronaux et des graphes pour un meilleur diagnostic des troubles cérébraux en utilisant des images TEP au 18F-FDG: Un pas vers un diagnostic personnalisé

Discipline
Biologie santé

Spécialité
Neurosciences

École doctorale
SCIENCES DE LA VIE ET DE LA SANTE (ED62)

Laboratoire de l'Institut Fresnel
Equipe IMOTHEP

Composition du jury

Olivier HUMBERT Université Côte d'Azur	Rapporteur
Valérie LOUIS-DORR Université de Nancy	Rapporteure
Régine LE BOUQUIN JEANNÉS Université de Rennes 1	Examinatrice
Nguyen Linh Trung Vietnam National University	Examinateur
Régine LE BOUQUIN JEANNÉS Université de Rennes 1	Présidente du jury
Eric GUEDJ Aix-Marseille Université	Directeur de thèse
Mouloud ADEL Aix-Marseille Université	Co-directeur de thèse

Institute Marseille Imaging Aix*Marseille Université
INSTITUT FRESNEL
CERIMED
FRANCE 2030

Affidavit

I, undersigned, Pham Minh Tuan, hereby declare that the work presented in this manuscript is my own work, carried out under the scientific supervision of Prof. Eric GUEDJ, Prof. Mouloud ADEL, and Prof. Nguyen Linh Trung in accordance with the principles of honesty, integrity and responsibility inherent to the research mission. The research work and the writing of this manuscript have been carried out in compliance with both the french national charter for Research Integrity and the Aix-Marseille University charter on the fight against plagiarism.

This work has not been submitted previously either in this country or in aother country in the same or in a similar version to any other examination body.

Marseille, le 20 Avril 2024

Pham Minh Tuan



This work is licensed under Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International Public License

Liste des publications et participations aux conférences

Liste des publications :

1. P. M. Tuan, T.-L. Phan, M. Adel, E. Guedj and N. L. Trung," AutoEncoder-based feature ranking for Alzheimer Disease classification using PET image," Machine Learning with Applications, vol. 6, p. 100184, 12 2021.

Participation aux conférences:

1. P. M. Tuan, T. -L. Phan, M. Adel, S. Bourennane, N. L. Trung and E. Guedj, "C-Atlas: A Brain Mapping based on FDG-PET Images for Alzheimer's Disease Diagnosis," 2022 RIVF International Conference on Computing and Communication Technologies (RIVF), Ho Chi Minh City, Vietnam, 2022, pp. 150-155, doi: 10.1109/RIVF55975.2022.10013857.

2. P. M. Tuan, N. L. Trung, M. Adel and E. Guedj, "AutoEncoder-based Feature Ranking for Predicting Mild Cognitive Impairment Conversion using FDG-PET Images," 2023 IEEE Statistical Signal Processing Workshop (SSP), Hanoi, Vietnam, 2023, pp. 720-724, doi: 10.1109/SSP53291.2023.10208072.

3. P. M. Tuan, M. Adel, N. L. Trung and E. Guedj, " Dynamic Time Warping-based distance for constructing individual brain networks from FDG-PET Images: Application to Alzheimer's Disease," 2023 IEEE Image Processing Theory, Tools and Applications (IPTA), Paris, France, 2023.

4. P. M. Tuan, M. Adel, E. Guedj and N. L. Trung, "Wasserstein-based distance for multi-scale individual brain network" submitted to The European Signal Processing Conference 2024 (EUSIPCO).

5. P. M. Tuan, M. Adel, E. Guedj and N. L. Trung, "Enhancing Depression Diagnosis Using FDG-PET Images with HyperGraphs" submitted to The European Signal Processing Conference 2024 (EUSIPCO).

Résumé

Le cerveau, en tant que partie du système nerveux, contrôle nos pensées et toutes les activités à travers notre corps. Avec environ 86 milliards de neurones connectés par 150 billions de synapses, le cerveau humain est l'un des réseaux les plus complexes. Les perturbations dans les réseaux cérébraux jouent un rôle important dans de nombreux troubles psychiatriques et neurologiques, tels que la dépression, la maladie de Parkinson et la maladie d'Alzheimer, pour n'en citer que quelques-uns. Différents types de troubles cérébraux peuvent présenter plusieurs similitudes en termes de symptômes cliniques et d'anomalies des réseaux cérébraux, ce qui rend difficile leur distinction. Il est donc crucial de mieux comprendre et de clarifier les différences de pathologies entre ces troubles cérébraux, et de mieux identifier les groupes anormaux par rapport au groupe témoin normal. Pendant ce temps, les biomarqueurs d'imagerie par tomographie par émission de positons (TEP) offrent de grandes promesses en tant que marqueurs diagnostiques et pronostiques pour les troubles cérébraux.

Dans ce but, cette thèse explore des approches innovantes pour améliorer le diagnostic de la maladie d'Alzheimer (MA) en utilisant des images FDG-PET et l'apprentissage automatique, en se concentrant particulièrement sur une méthodologie basée sur les régions, avec des implications pour l'application dans les troubles psychiatriques également. Elle aborde les principaux défis du diagnostic de la MA, tels que la sélection des caractéristiques, le choix de l'atlas cérébral et l'importance de l'analyse à plusieurs échelles. De plus, elle explore également la méthode de construction de réseaux cérébraux individuels métaboliques. En mettant l'accent sur l'importance des réseaux neuronaux graphiques et des hypergraphes, la recherche étend son application au-delà des troubles neurologiques tels que la MA pour englober les troubles psychiatriques tels que la dépression. Plus précisément, quatre travaux indépendants ont été menés.

Le premier travail exploite l'Autoencodeur (AE), un type spécifique de réseau neuronal, pour classer et sélectionner les régions cérébrales les plus importantes lors de la cartographie de l'évolution de la MA. Au lieu de se concentrer sur des architectures AE complexes pour la sélection des caractéristiques, nous proposons le concept de "AE unique mais multiple" pour les caractéristiques de sélection, et montrons son efficacité dans le classement des régions cérébrales affectées par la MA ainsi que dans l'amélioration des performances globales de diagnostic.

Le deuxième travail examine l'efficacité de la cartographie cérébrale adaptée et de la cartographie cérébrale à plusieurs échelles pour le diagnostic de la MA. Nous proposons une méthode de construction de la cartographie cérébrale adaptée et à

plusieurs échelles appelée C-Atlas pour un meilleur diagnostic de la MA.

Le troisième travail propose des méthodes de construction de réseaux cérébraux individuels et de graphiques en tant qu'outils efficaces pour l'analyse des troubles cérébraux. Ces méthodes ont montré leur efficacité dans la capture des changements entre les différentes étapes de la MA, et l'efficacité des graphiques pour modéliser les changements dans la connectivité des troubles cérébraux.

Le dernier travail se concentre sur les réseaux neuronaux graphiques et les hypergraphes comme outils meilleurs non seulement pour diagnostiquer la MA mais aussi pour les troubles psychiatriques tels que la dépression. Nous avons mené des expériences avec différents types courants de réseaux neuronaux graphiques, afin de mettre en évidence la contribution des réseaux neuronaux graphiques à un meilleur diagnostic.

Les expériences menées sur des ensembles de données locaux et publics montrent que nos études non seulement font progresser le diagnostic de la MA, mais fournissent également des indications sur la contribution des réseaux neuronaux et des réseaux neuronaux graphiques à un meilleur diagnostic des troubles cérébraux, ouvrant la voie à des méthodologies de diagnostic améliorées et à des résultats améliorés pour les patients.

Mots clés: TEP, troubles neurologiques, troubles psychiatriques, maladie d'Alzheimer, dépression, réseaux cérébraux, réseaux individuels, autoencodeur, cartographie cérébrale, réseaux neuronaux graphiques.

Abstract

The brain, as part of the nervous system, controls our thoughts and all activities throughout our body. With approximately 86 billion neurons connected through 150 trillion synapses, the human brain is one of the most complex networks. Disruptions in brain networks play an important role in many psychiatric and neurological disorders, such as depression, Parkinson's disease, and Alzheimer's disease, to name a few. Different types of brain disorders may have several similarities in terms of clinical symptoms and brain network abnormalities, making it difficult to distinguish between them. Therefore, it is crucially important to better understand and clarify the differences in pathologies between those brain disorders, and better identifying the abonormal groups from the normal control group. Meanwhile, Positron Emission Tomography (PET) imaging biomarkers hold great promise as diagnostic and prognostic markers for brain disorders.

For this purpose, this thesis delves into innovative approaches to enhance Alzheimer's disease (AD) diagnosis using FDG-PET images and Machine Learning, particularly focusing on a region-based methodology, with implications for application in psychiatric disorders as well. It addresses key challenges in AD diagnosis, such as feature selection, brain atlas choices, and the importance of multiple-scale analysis. Additionally, it also explores the method for building metabolic individual brain networks. Emphasizing the significance of graph neural networks and hypergraphs, the research extends its application beyond neurological disorders like AD to encompass psychiatric disorders such as depression. Specifically, four independent works have been conducted.

The first work leverages Autoencoder (AE), a specific type of neural network, for ranking and selecting the most important brain regions during the course map of AD. Instead of focusing on complex AE architectures for feature selection, we propose the concept of "single-but-multiple" AEs for feature selection features, and show its effectiveness in ranking brain regions affected by AD as well as improving the overall diagnosis performance.

The second work investigates the effectiveness of adapted brain mapping and multiple-scale brain mapping for AD diagnosis. We propose a method to build adapted and multiple-scale brain mapping called C-Atlas for better diagnosis of AD.

The third work proposes methods for building individual brain networks and graphs as effective tools for analyzing brain disorders. These methods have shown effectiveness in capturing changes between various stages of AD, and the effectiveness of graphs for modeling the changes in connectivity of brain disorders.

The last work focuses on graph neural networks and hypergraphs as better tools not only for diagnosing AD but also for psychiatric disorders such as depression. We conducted experiments with different common types of graph neural network such, in order to emphasize the contribution of graph neural networks to better diagnosis.

Experiments conducted on both local and public datasets show that our studies not only advance AD diagnosis but also provide insights into the contribution of neural network and graph neural networks toward better diagnosing brain disorders, paving the way for enhanced diagnostic methodologies and improved patient outcomes.

Keywords: FDG-PET images, neurological disorders, psychiatric disorders, Alzheimer's disease, Depression, brain networks, individual networks, autoencoder, brain mapping, graph neural networks.

Acknowledgments

This thesis marks the end of my three-year journey towards a Ph.D. degree in Neuronscience from the Aix-Marseille University. Despite having some difficulties and things that didn't go as expected, I have still enjoyed being a doctoral student at the Fresnel Insitute, CERIMED, Aix-Marseille University. I am so grateful to have had the opportunity to meet, work with, and learn from so many wonderful people.

To my supervisors,

First of all, I would like to express my great appreciation to Prof. Eric GUEDJ and Prof. Mouloud ADEL, my supervisors at the Aix-Marseille University, and Prof. Nguyen Linh Trung, my advisor at VNU University of Engineering and Technology. I am lucky to work under the supervision of prof. Eric and Prof. Mouloud who are very kind, friendly, and unconditionally supportive throughout my study. Thank you very much for giving me a lot of freedom to pursue my ideas in research, and for all useful comments on my works in regular meetings. Without the support of Prof. Trung, I wouldn't have got where I am today, thank you so much.

To my family,

To my parents and my little brother. I am immensely thankful to my family for their love, care, and extremely supports. I love them from my heart as always happens.

To my colleagues and my friends,

I am very fortunate to have a lot fantastic friends: Thanh, Tuan, Ha, Dinh, Ngoc, Linh whose sincerity and friendship are highly appreciated. My Ph.D. journey has been much more fun and exciting with their stories, talks, and moments. My life in France would not have been as pleasant as it is right now without their help.

Last but not least, thanks to Excellence Initiative of Aix-MarseilleUniversity - A*MIDEX (program AMX-19-IET-002) and Marseille Imaging Insitute for their funding, supported during the three years.

Thank you!
Pham Minh Tuan

Contents

Affidavit 2

Publications 3

Résumé 4

Abstract 6

Acknowledgments 8

Contents 9

List of Figures 13

List of Tables 15

Préntation du travail en Français 17

1 Introduction 19
- 1.1 Context and Problem Statements . . . 19
- 1.2 Scope and Objectives of thesis . . . 22
- 1.3 Contributions and outline . . . 24

2 Backgrounds 26
- 2.1 Introduction to Alzheimer's Disease . . . 27
  - 2.1.1 Alzheimer's Disease . . . 27
  - 2.1.2 Alzheimer's Disease Hypothesis . . . 28
  - 2.1.3 Progression of Alzheimer's Disease . . . 29
  - 2.1.4 Alzheimer's Disease Diagnosis . . . 30
- 2.2 Introduction to NeuroImaging . . . 32
  - 2.2.1 (f)MRI and its use . . . 32
  - 2.2.2 PET and its use . . . 34
  - 2.2.3 SPECT . . . 36
  - 2.2.4 CT and its use . . . 37
  - 2.2.5 M/EEG and its use . . . 38
  - 2.2.6 Other modalities . . . 38
- 2.3 Machine Learning . . . 38
  - 2.3.1 Machine Learning pipeline . . . 38
  - 2.3.2 Features . . . 39

2.3.3 Dimensionality reduction . . . 40
2.3.4 Learning approaches . . . 41
2.3.5 Validation . . . 42
2.4 ADNI dataset . . . 45
2.4.1 Introduction . . . 45
2.4.2 Data acquisition . . . 46
2.4.3 Data selection . . . 47
2.4.4 Data processing . . . 48
2.5 The state of the art . . . 48

3 Autoencoder-based feature ranking 50
3.1 Introduction . . . 51
3.2 Autoencoder-based feature ranking for AD diagnosis . . . 52
3.2.1 Related work . . . 53
3.2.2 Proposed method . . . 55
3.2.3 Results and Discussion . . . 60
3.2.4 Conclusion . . . 74
3.3 Autoencoder-based feature ranking for MCI diagnosis . . . 75
3.3.1 Proposed Method . . . 76
3.3.2 Results and Discussion . . . 79
3.3.3 Conclusions and Future work . . . 87
3.4 Chapter Conclusion . . . 87

4 Adapted Brain Atlas 89
4.1 Introduction . . . 89
4.2 C-Atlas . . . 90
4.2.1 Introduction . . . 90
4.2.2 Proposed method . . . 91
4.2.3 Proposed method . . . 93
4.2.4 RESULTS AND DISCUSSION . . . 95
4.2.5 CONCLUSIONS AND FUTURE WORK . . . 100
4.3 Chapter Conclusion . . . 101

5 Metabolic Network: A Short Review 102
5.1 Introduction . . . 104
5.2 Methodology . . . 105
5.2.1 Research Questions . . . 105
5.2.2 Search Strategy and Eligibility Criteria . . . 106
5.3 Basic Concepts . . . 106
5.3.1 Graph Theory . . . 106
5.3.2 Construction of Metabolic Connectivity network using FDG-PET 107
5.3.3 Computing Graph Measures . . . 108
5.3.4 Most common graph metrics . . . 109

5.4 Results . . . 112
5.4.1 Literature Search . . . 112
5.4.2 Study Characteristics . . . 113
5.4.3 General Overview . . . 113
5.5 Discussion . . . 117
5.5.1 Healthy Control-based Metabolic Network Analysis . . . 117
5.5.2 Metabolic Network Patterns in Brain Disorders . . . 118
5.6 Challenges and Future Directions . . . 123
5.6.1 From Data Perspective . . . 123
5.6.2 Building and Analyzing brain metabolic networks . . . 123
5.6.3 Neural Network as a tool for exploring brain connectivity . . . 125
5.6.4 Clinical Intervention . . . 125
5.7 Conclusion . . . 125
5.8 Chapter Conclusion . . . 126

6 Metabolic Individual Brain Network Analysis with Graph theory 128
6.1 Introduction . . . 129
6.2 Materials and Methods . . . 130
6.2.1 Dataset . . . 130
6.2.2 Methods . . . 131
6.3 Results and Discussion . . . 133
6.3.1 Results . . . 133
6.3.2 Discussion . . . 134
6.4 Conclusions . . . 136
6.5 Chapter Conclusion . . . 136

7 Graph Neural Network Analysis 137
7.1 Introduction . . . 138
7.2 Background . . . 139
7.2.1 Related works . . . 139
7.2.2 Graph Notations and Graph Building . . . 140
7.2.3 Graph Neural Networks for Graph Classification . . . 140
7.2.4 Graph Readout . . . 142
7.3 Data & Experimental Setup . . . 142
7.3.1 Datasets . . . 142
7.3.2 Graph Building . . . 144
7.3.3 Model selection & assessment . . . 145
7.3.4 Non-graph baselines . . . 145
7.4 Results . . . 145
7.4.1 Disorder diagnosis task (dataset) . . . 145
7.4.2 Network Scale . . . 148
7.4.3 Network building method . . . 148
7.4.4 Graph Neural Network Models . . . 151

7.4.5 Interplay between network scales, network building methods and graph neural network models . . . 151
7.5 Discussion . . . 152
7.6 Conclusion . . . 153

8 Enhancing Depression Diagnosis Using FDG-PET Images with HyperGraphs 154
8.1 Introduction . . . 155
8.2 Materials and Methods . . . 156
8.2.1 Datasets . . . 156
8.2.2 Methods . . . 157
8.2.3 Experimental settings . . . 159
8.3 Results and Discussion . . . 160
8.3.1 Results . . . 160
8.3.2 Discussion . . . 161
8.4 Conclusion . . . 162

9 Conclusion 163

Conclusion 163

Bibliography 166

List of Figures

1.1 Biomarker changes in the progression of AD 22

2.1 Amyloid Plaques and Neurofibrillary Tangles. 28
2.2 Evolution of AD in the brain. 29
2.3 Two representative MRI images from the NC and AD groups. 32
2.4 Differences of functional connectivity between NC and AD captured by rs-fMRI (axial view) 33
2.5 Diffusion tensor MRI in Alzheimer's research. 34
2.6 FDG-PET of control, MCI, and mild AD individuals. 35
2.7 PIB-PET of control, MCI, and mild AD individuals. 36
2.8 Differences between NC and AD captured by SPECT 37
2.9 Support Vector Machine (SVM) 41
2.10 An instance of ROC curve. 44
2.11 An illustration of 5-fold cross-validation. 45

3.1 The flowchart of the proposed method. 55
3.2 Simple Autoencoder structure. 58
3.3 Fine-Tuning AutoEncoder. 58
3.4 The flow chart of 2-phase feature ranking method. 59
3.5 Performance of 4 baseline methods with different numbers of ROIs. . . 62
3.6 Comparing performance between proposed method and the other methods under different numbers of ROIs. 63
3.7 The top-20 ROIs selected by proposed method on 3D brain image. . . 63
3.8 Comparing performance between 1st phase of proposed method and the other methods under different numbers of ROIs. 64
3.9 Comparing performance between 1st phase of proposed and Lasso method when the number of selected features is in range of [9,54]. . . 65
3.10 Comparing performance between the first and the second phase of proposed method under different numbers of ROIs. 66
3.11 Comparing position Index of the all ROIs before and after the 2nd phase. 67
3.12 Flow chart of the proposed method 76
3.13 Comparing ACC of linear SVM classifier with and without ranking methods. 81
3.14 Comparing ACC of linear SVM classifier between our proposed methods and baselines. 82
3.15 Comparing AUC of linear SVM classifier between our proposed methods and baselines. 82

3.16 Overlapped rate between different ranking methods. 83
3.17 An illustration for top-20 ROIs selected by our proposed method (with fixing connections) 87

4.1 The flowchart of the proposed method, including 2 phases: SVM-based (top) and SLIC-based (bottom). 93
4.2 Performance comparison between C-Atlas and Schaefer. 97
4.3 Boxplot for top-9 ROIs of C-Atlas (200), comparing distribution of data between NC and AD group. 98
4.4 Comparison visualization between AAL (120 regions) and C-Atlas (200 regions). 99

5.1 Publication trend by year of keyword-matched studies (left) and included studies (right). 113
5.2 Categorization of included studies based on data group. 114
5.3 Word Clouds illustrate the top 50 of most frequency words in title (left) and abstract (right) of keyword-matched articles. 115
5.4 Word Clouds illustrate the top 50 of most frequency words in title (left) and abstract (right) of included articles. 115
5.5 Collocation Analysis for words in Abstracts of included articles. 116

6.1 Workflow for constructing an individual graph from a static FDG-PET image. 130
6.2 Comparing individual graph metrics between different groups 134
6.3 Comparing individual graph metrics between different groups 134
6.4 Comparing individual graph metrics between different groups 135

7.1 Pipeline of experiments. 143
7.2 Comparison of average performance of baseline methods (MLP, RBF SVM) with GNNs (GCN, GIN, GAT) on ADNI and psychiatric datasets across various metrics. 146
7.3 Performance comparison of all methods across different scales (100, 200, 300 regions) on the psychiatric dataset. 146
7.4 Performance comparison of all methods across different scales (100, 200, 300 regions) on the ADNI dataset. 146
7.5 Performance comparison of all methods when comparing two graph building methods (DTW, KLS) on the psychiatric dataset. 147
7.6 Performance comparison of all methods when comparing two graph building methods (DTW, KLS) on the ADNI dataset. 147
7.7 Performance comparison of all methods when graph neural network structure (GCN, GIN, GAT) on the psychiatric dataset. 147
7.8 Performance comparison of all methods when graph neural network structure (GCN, GIN, GAT) on the ADNI dataset. 148

8.1 Pipeline of experiments. 157

List of Tables

2.1 Confusion Matrix for a Binary Classification 43
2.2 Demographic and Clinical information of subjects 48
2.3 Performance of SVM Models for AD Detection 49

3.1 Demographic and clinical information of subject 57
3.2 Performance of 4 baseline ranking methods 68
3.3 Overlap rate when comparing 4 baseline ranking methods 69
3.4 Shared rois between lasso, fisher, t-test when selecting top-20 rois . . . 70
3.5 Comparing accuracy of SVM classifiers between proposed method with baseline methods 70
3.6 Detailed performance of SVM classifiers for the proposed method . . . 71
3.7 Overlap rate when comparing proposed method with baseline methods 71
3.8 Top 20 rois selected by proposed method 72
3.9 DEMOGRAPHIC AND CLINICAL INFORMATION OF SUBJECT 77
3.10 Performance on sMCI/pMCI classification task 80
3.11 Top 20 rois selected by proposed method 86

4.1 Demographic and clinical information of subject 92
4.2 Performance on AD/NC classification task 96
4.3 Location of C-ATLAS (200) Top-Ranked Regions comparing to AAL . . . 100

5.1 List of included studies in this survey 127

6.1 DEMOGRAPHIC AND CLINICAL INFORMATION OF SUBJECT 130
6.2 Classification performance on AD classification tasks. 133
6.3 Comparing Graph metrics between different groups 133

7.1 DEMOGRAPHIC AND CLINICAL INFORMATION OF SUBJECT 144
7.2 Comparison of Performance Metrics for various models on ADNI and Psychiatric Local Dataset with a network proportion of 0.2. 149
7.3 Comparison of Performance Metrics for various models on ADNI and Psychiatric Local Dataset with a network proportion of 0.5. 150

8.1 Comparison of Performance Metrics for GNN models with pairwise graphs. 160
8.2 Comparison of Performance Metrics for GNN models with hypergraphs. 160

Préntation du travail en Français

Cette thèse aborde les problèmes liés au diagnostic de la maladie d'Alzheimer (AD) en utilisant des images FDG-PET et l'apprentissage automatique, en se concentrant sur une approche basée sur la région. Elle présente des méthodes pour étudier plusieurs aspects cruciaux du diagnostic de la maladie d'Alzheimer :

- Le chapitre 1 introduit le contexte scientifique, les objectifs, le plan et les contributions de cette thèse.
- Le chapitre 2 fournit des informations de base sur la maladie d'Alzheimer (AD), y compris les facteurs cruciaux, l'évolution de l'AD et les outils de diagnostic. Il couvre également l'imagerie cérébrale, en se concentrant particulièrement sur les images FDG-PET, et les méthodes de diagnostic assisté par ordinateur pour l'AD, englobant l'extraction de caractéristiques, la sélection de caractéristiques, l'apprentissage automatique, l'apprentissage profond (avec un accent sur les autoencodeurs), les classificateurs, l'évaluation des performances et l'état actuel de l'art en termes de performance. De plus, il aborde l'ensemble de données ADNI.
- Le chapitre 3 présente une méthode de classement des caractéristiques basée sur l'autoencodeur pour sélectionner des caractéristiques significatives à différents stades de la maladie d'Alzheimer, préconisant l'utilisation de plusieurs modèles simples plutôt que d'architectures complexes.
- Le chapitre 4 propose une approche basée sur les données utilisant des images FDG-PET pour construire une cartographie cérébrale adaptative pour le diagnostic de la maladie d'Alzheimer, en mettant l'accent sur l'importance du choix de l'atlas cérébral et de l'analyse à plusieurs échelles.
- Le chapitre 5 fournit un aperçu de l'analyse graphique et métabolique, mettant en évidence leur rôle dans la compréhension des interactions entre plusieurs régions cérébrales.
- Le chapitre 6 présente des méthodes pour construire des réseaux cérébraux individuels, démontrant leur efficacité dans la capture des changements liés à la maladie d'Alzheimer et des variations individuelles.
- Le chapitre 7 réalise une analyse comparative des architectures de réseaux neuronaux graphiques pour la classification des troubles neurologiques et psychologiques, mettant en avant leur contribution à l'apprentissage de bout en bout.

- Le chapitre 8 examine l'efficacité des hypergraphes dans la capture de schémas complexes dans les troubles cérébraux des patients, en particulier dans les cas de différences subtiles entre les groupes de patients et de contrôle.

Dans l'ensemble, cette thèse contribue à faire progresser le diagnostic de la maladie d'Alzheimer en intégrant des techniques d'apprentissage automatique, des méthodes de cartographie cérébrale et des approches d'analyse graphique, ouvrant la voie à une meilleure compréhension et à un meilleur diagnostic des troubles cérébraux.

Chapter 1

Introduction

Summary

1.1 Context and Problem Statements 19
1.2 Scope and Objectives of thesis 22
1.3 Contributions and outline 24

1.1 Context and Problem Statements

Alzheimer's disease (AD) is the most prevalent form of brain disorders in ageing population, which accounts for 50% to 70% of cases [168]. The most common early symptom is loss of short-term memory, and as the disease progresses, different symptoms may gradually appear, including language disorders, disorientation, emotional instability, loss of motivation and many behavioural problems [70, 205, 247]. With the situation worsening, patients tend to gradually lose their physical function, eventually leading to death [70]. The World Alzheimer Report 2015 estimates there are approximately 50 million people living with dementia worldwide [168] and the number of individuals with AD is expected to double by 2030 and to reach 115 million by 2050 [73]. Alzheimer's disease affects one in nine over 65s and one in three over 85s [88]. As we are getting older, AD is unavoidable and the health care costs associated with AD are exceptionally high, imposing an important burden on modern societies. At present, AD is incurable, even if promising developments for treatments are expected to be achieved in the near future.

AD is a chronic neurodegenerative disorder that leads to progressive decline of cognitive functions along with behavioural disorders and insidious loss of autonomy in daily living activities. The evolution of AD follows five stages. The "preclinical" stage is asymptomatic, but the brain lesions of AD are present. At this stage, the concept of Subjective Cognitive Impairment (SCI) has been proposed recently, defined by a self-experienced persistent decline in cognitive capacity in comparison with a previously normal status [69]. These subjective complaints are considered as a risk factor for AD [168, 108]. Then, in the "prodromal" stage, Mild Cognitive Impairment (MCI) patients exhibit measurable memory impairments, but maintain their functional capacities

[108, 140]. In the "Mild AD" stage, cognitive deficits are more notable, such as memory and learning impairments. These symptoms become more severe in the "Moderate AD" stage. In the final "severe" stage of the disease, almost all cognitive and motor functions are deeply deteriorated and patients lose autonomy becoming completely dependent on caregivers. The average duration of survival of AD patients is 5-8 years after clinical diagnosis [63]. Nowadays, no medication exists for curing this pathology, and many therapeutics trials failed. Current major anti-AD medication trials are focusing on the preclinical stage, in order to treat before symptoms development. Therefore, AD detection at MCI and preclinical stages becomes an important issue for the scientific community. It allows to either allow patients to receive interventions at an early stage or provide insights into the disease progression.

The most commonly used criteria for diagnosis of Alzheimer's disease are provided by the National Institute of Neurological and Communicative Disorders and Stroke - Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) [136, 197, 7]. Accordingly, AD is usually diagnosed based on the person's medical history, history from relatives, and clinical examination. Neuropsychological tests can further characterize the state of the disease, including a memory test such as California Verbal Learning Test (CVLT) [125] or some comprehensive tests which combine a range of tests to provide an overview of cognitive skills, like mini–mental state examination (MMSE)[141] and clinical dementia rating sum of boxes (CDR-SB) [130] etc. The NINCDS-ADRDA criterion has also suggested the utility of different biomarkers of the pathophysiological process to weight the diagnostic probability of the disease, such as cerebrospinal fluid (CSF) biomarkers (Aβ42, t-tau, p-tau), blood biomarkers, genetic biomarkers (APOE ϵ4), as well as neuroimaging modalities (Positron Emission Tomography - PET, functional Magnetic Resonance Imaging fMRI, Single Photon Emission Computed Tomography - SPECT). Fig. 1.1 shows changes of different biomarkers during the progression to AD. In which, Aβ is measured by cerebrospinal fluid or by amyloid PET imaging. Neurodegeneration is determined by tau protein measured in cerebrospinal fluid, or by synaptic dysfunction, measured by FDG-PET. Brain atrophy, mostly in the medial temporal lobe, measured by structural MRI. Memory loss, measured by cognitive tests. Clinical function, indicated by general cognitive decline measured by cognitive tests. As can be seen, biomarker changes have occurred before clinical dysfunction. Amyloid beta accumulation identified by CSF or PET is the most sensitive, followed by synaptic dysfunction.

Recently, there were growing interesting in using neuroimaging modalities as biomarkers for AD. Among those kinds of neuroimaging biomarkers, functional neuroimaging can be considered as one of the most effective neuroimaging biomarkers. Comparing to structure neuroimaging biomarkers, functional biomarkers can capture well and early changes in brain functions, which are so useful for diagnosing at earlier stages of AD. Among those functional neuroimaging, neuroimaging such as Fluorodeoxyglucose PET (FDG-PET) is suggested by NINCDS-ADRDA to use in clinical as requirements steps for diagnosing AD. Event though, due to the fact that the complex nature of imaging, and it is time consuming for diagnosing AD, there is a growing interest in using Machine learning and computer aided diagnosis methods for better

diagnosis. Each year, there are more than hundreds of studies regarding this research topics. And it is expected that this number will increasing more in the future. ***This thesis investigates the diagnosis of AD at various stages using FDG-PET images and machine learning approach. More specifically, in this work, we considered several machine learning approach and graph as tools for better analyzing FDG-PET images and better diagnosis AD using FDG-PET images.***

One of the main problems of using FDG-PET images in particular, and neuroimaging in general for diagnosing AD using machine learning is "dimensional-curve". The dimension of data can be vary high, it's make challenges for a classifier for a better performance. In order to overcome this limitation, there are many techniques have been proposed recently year. One of the most common way is to use brain mapping as a tool for separating brain image into separated areas, based on that features are computed and used for classifiers. This approach we called it as "region-based" or region-of-interests (ROIs) analysis using brain mapping. This approach show several advantages compared to other approaches as it help reducing dimension of data, in the same time, by using predefined mapping, we can mapping changes in brain with brain functions, and reasoning about types of changes, meaning of changes.

Eventhough, there is a needs for determined which are the most important brain regions and the interactions between them. Considering that, during various stages of AD, some regions may be more important at this steps, while the other maybe important at later phases, while there are regions may very stables or continuously changes during various stages of AD. Determined these regions are extremely important for tracking changes of disease as well as effective treatments. Moreover, even when affected in the same stage of AD, the magnitude of each region's effect may vary greatly from one another. In the first part of this thesis, we addressed these questions by proposing methods based on Autoencoder to select the most affected brain regions by AD and provide rankings for each region based on their contribution to the performance of classifier diagnosis.

Considering that the choices of brain mapping is extremely important and may has a lot of effects to final analysis results. Meanwhile, predefined atlas, in general, are for general studies about brain, which may not be suitable for diagnosis a specific brain disorders – due to the fact that brain may be divided into regions maybe too small or too large to represent for a brain disorders, in the second part of this thesis we proposed to build an adaptive atlas for better diagnosis of AD, named C-Atlas, this atlas is expected to improve a better diagnosis performance of brain disorders.

Considering that several regions may change or be affected in the same stages during the course of AD, it's essential to account for the change patterns between multiple brain regions and their interactions. Therefore, in the third part of this thesis, we utilized graph analysis as a tool for better understanding and diagnosing AD. Specifically, we proposed a method to construct individual brain networks from FDG-PET images. These initiatives aim to take a step toward personalized diagnosis.

The complexity of relationships within brain graphs poses a significant challenge for accurate diagnosis. To address this complexity, the utilization of graph neural networks emerges as a powerful tool capable of automatically discerning intricate pat-

terns inherent in brain graphs. By comparing various building settings, our research underscores the pivotal role of graph neural networks in enhancing the diagnostic process for brain disorders. Through their ability to capture nuanced interconnections within brain networks, graph neural networks offer a promising avenue for achieving more precise and reliable diagnoses, thereby advancing our understanding and treatment of neurological conditions.

In our efforts to improve diagnostic accuracy, especially in psychiatric disorders like depression, where conventional methods struggle with subtle changes, we explore hypergraph and graph neural network technologies. These advanced tools excel in capturing complex interactions within brain networks, surpassing the limitations of traditional graphs. By utilizing hypergraph and graph neural network frameworks, we aim to enhance diagnostic capabilities.

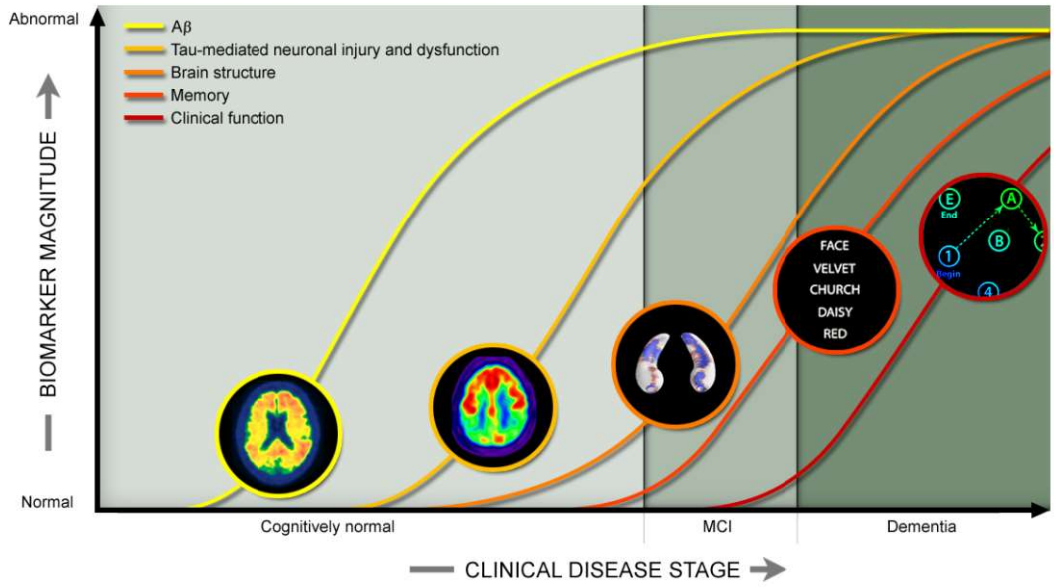

Figure 1.1: Biomarker changes in the progression of AD https://adni.loni.usc.edu/study-design/.

1.2 Scope and Objectives of thesis

In this thesis, we focus on 3 research subjects:

1. *FDG-PET images* as Input
2. *Machine learning (Autoencoder, Graph, Graph Neural Networks)* as study methods
3. *Computer Aided Diagnosis of brain disorders*: Mainly focused on improving the diagnosis of Alzheimer's disease at various stages. Based on this focus,

proposing several enhancements for the analysis of psychiatric disorders, such as depression, where the contrasts between control and patient groups are more subtle.

We focus on region-based approach, which consider to dividing brain into areas and study. The problem we focus on is study determining the most important brain regions affected by AD, ranking their important during various stages, ranking (autoencoder-based feature ranking); the choices of brain mapping (C-Atlas); the interactions between different brain regions (individual graph, graph neural network, hypergraph). All of this works toward a better diagnosis of AD, based on that moving to applications in psychiatric disorders. Each work answer a separate question, however, when putting together, they give a better picture regarding to region-based approach, as well as providing evidence for the effective of ML-based approach as tool for better diagnosis of brain disorders. Specifically, in this thesis, we study the following questions:

1. **Which is the most important during the course map of AD? (Are there any stable regions, which one changes at various stages, which one can be used as markers for transit between various stages of AD)?** In order to answer this question, we use Autoencoder as methods for selecting the most important regions. Specifically, instead of designing a complex architecture, we propose to use simple-but-multiple AEs for selecting the most important features. To this end, we proposed Autoencoder-based features ranking frameworks, this highlight the advantages of "simple-but-multiple" AEs. As well as highlight the importance of several brain regions during the course map of AD.

2. **Is the choice of brain mapping important, could we use adapted atlases learning from data for AD diagnosis?** In order to answer this question, we proposed a method to build an adaptive atlas from data, named C-Atlas. This atlas has show several advantages compared to traditional brain atlas. It's not providing way to separate brain into regions, but also give the rank – determined how much important of brain regions to AD.

3. **Are there any region that change at the same time? Or do they change in same velocity?** In order to study that, we use graph as mathematical frameworks for study the changes and interaction between various brain regions during the course maps of AD. Specifically, we proposed a method for building individual brain networks, it's means that from each static FDG-PET image, we construct a graph which represents for the interactions between various regions. And we use these graph study the changes patterns in brain networks.

4. **Lastly, the contribution of graph neural network to enhancing diagnosis in neurological and psychiatric disorders?** In our investigation, we undertake empirical analyses involving several prevalent graph neural network architectures, focusing on their application in AD and Depression diagnosis. Through rigorous examination, we evaluate the impact of various building methods, network architectures, and training settings on the classifiers' performance. Additionally, we

delve into the potential of hypergraphs as a superior framework for augmenting the efficacy of graph neural networks.

1.3 Contributions and outline

The novelty of our works lie in the two points:

1. **Group-based diagnosis, the most important brain regions:** We show an example about the effective of using simple-but-multiple AE structures for selecting the most important features. It suggests that AE not only used for feature extraction but can also be used for feature selections. We proposed a concepts of adapted brain mapping for a specific brain disorders. This works, also help to selected the most important course map of AD.

2. **Individual-based diagnosis, the interaction between brain regions:** We studies two methods for individual brain networks and study changes and interactions between multiple regions during the course map of AD.

3. **Graph Neural Networks:** Our study delves deeply into the efficacy of employing graph neural networks and hypergraphs as advanced tools for probing metabolic brain networks. We showcase their potential contributions in both neurological and psychological disorder analysis, highlighting the innovative approaches they offer in understanding these complex systems.

Collectively, These works advance our understanding and diagnosis of Alzheimer's disease through Computer-Aided Diagnosis. Shifting from group-based to individual-based diagnostics is crucial for personalized healthcare. Additionally, they highlight the importance of graph neural networks in diagnosing brain disorders and emphasize the unique value of hypergraphs, especially in conditions like depression, where subtle changes are significant.

This thesis comprises 9 chapters:

- Chapter 1 introduces the scientific context, objectives, outline, and contributions of this thesis.
- Chapter 2 introduces the background on:
  - Alzheimer's disease (AD), crucial factors in AD, the evolution of AD, and diagnostic tools for AD.
  - Neuroimaging, particularly focusing on FDG-PET images.
  - Computer-aided diagnosis methods for AD, covering:
    - Feature Extraction, Feature Selection.
    - Machine Learning and Deep Learning (focusing on AE).
    - Classifiers and Performance Evaluation.

    - The current state of the art (performances).
  - ADNI dataset.

- Chapter 3 covers Autoencoder-based approaches for AD and MCI diagnosis.
- Chapter 4 describes a data-driven brain mapping for AD diagnosis, C-Atlas.
- Chapter 5 introduces the concept of graphs and the analysis of metabolic brain networks.
- Chapter 6 investigates individual brain networks using Dynamic Time Warping (DTW) method and graph metrics.
- Chapter 7 investigates individual brain networks using graph neural networks.
- Chapter 8 hypergraph and graph neural network on psychiatric disorders, focusing on Depression.
- Chapter 9 provides conclusions and outlines future perspectives.

Chapter 2

Backgrounds

Summary

2.1 Introduction to Alzheimer's Disease . . . 27
2.1.1 Alzheimer's Disease . . . 27
2.1.2 Alzheimer's Disease Hypothesis . . . 28
2.1.3 Progression of Alzheimer's Disease . . . 29
2.1.4 Alzheimer's Disease Diagnosis . . . 30
2.2 Introduction to NeuroImaging . . . 32
2.2.1 (f)MRI and its use . . . 32
2.2.1.1 Anatomical MRI . . . 32
2.2.1.2 Functional MRI . . . 33
2.2.1.3 Diffusion MRI . . . 33
2.2.2 PET and its use . . . 34
2.2.2.1 FDG PET . . . 35
2.2.2.2 Amyloid PET . . . 35
2.2.3 SPECT . . . 36
2.2.4 CT and its use . . . 37
2.2.5 M/EEG and its use . . . 38
2.2.6 Other modalities . . . 38
2.3 Machine Learning . . . 38
2.3.1 Machine Learning pipeline . . . 38
2.3.2 Features . . . 39
2.3.3 Dimensionality reduction . . . 40
2.3.4 Learning approaches . . . 41
2.3.4.1 Support Vector Machine . . . 41
2.3.4.2 Logistic Regression . . . 42
2.3.4.3 Autoencoder . . . 42
2.3.5 Validation . . . 42
2.3.5.1 Performance metrics . . . 42
2.3.5.2 Cross-validation . . . 44
2.4 ADNI dataset . . . 45
2.4.1 Introduction . . . 45

2.4.2 Data acquisition 46
2.4.3 Data selection 47
2.4.4 Data processing 48
2.5 The state of the art 48

2.1 Introduction to Alzheimer's Disease

2.1.1 Alzheimer's Disease

AD is the most prevalent form of dementia, affecting a significant portion of the population aged over 65 years and accounting for most dementia cases [175, 68]. It is estimated that AD constitutes nearly 70% of all dementia cases. Furthermore, AD's incidence sharply increases with age, doubling in risk approximately every five years after reaching 65 [170, 47]. As individuals age, the risk of developing AD rises. Recent reports estimate a substantial increase in the likelihood of developing AD beyond the age of 75, estimated at around 17.8% [172]. It is anticipated that nearly 115 million individuals will be living with AD by 2050 [73]. This projection underscores the alarming trend of escalating AD cases, emphasizing its potential future impact on healthcare systems and society as a whole.

AD is characterized as a condition marked by progressive and irreversible brain damage, leading to cognitive dysfunctions such as impairments in thinking, reasoning, and memory. These dysfunctions trigger behavioral or personality changes in affected individuals. The onset of AD occurs well before the manifestation of dementia symptoms. As the disease advances, it causes severe memory impairments, resulting in a complete loss of autonomy. This loss significantly impacts an individual's ability to perform daily activities independently.

While the specific causes of AD remain not fully understood, the most common biological hypothesis suggests that AD involves the aggregation or accumulation of two proteins, amyloid-beta and tau. These proteins are often present in the brains of affected individuals during post-mortem examinations, indicating their potential role in the development and progression of AD.

Given the complex and not yet fully understood nature of AD, there is currently no medication available that can cure the disease. Despite numerous therapeutic trials, none have been successful in finding a definitive cure. Therefore, early identification of AD has become crucial. While existing treatments cannot halt or reverse the disease's progression, there is evidence to suggest that early interventions may help delay its advancement. This delay can be valuable, allowing patients to maintain their independence and functional abilities for a certain period. Additionally, early diagnosis of AD holds significance not only for patients but also for their caregivers. It enables both parties to prepare for the challenges that may arise in the future.

2.1.2 Alzheimer's Disease Hypothesis

The prevailing hypothesis about AD suggests that its clinical characteristics can be explained by the deposition of two specific types of structures: neuritic plaques and neurofibrillary tangles [181]. Neuritic plaques are abnormal clusters of proteins called amyloid-beta peptides (Aβ) that form outside neurons. These plaques can disrupt normal communication between neurons, impacting synaptic transmission [234, 52, 191]. On the other hand, neurofibrillary tangles are structures composed of hyperphosphorylated tau peptides and develop inside neurons (Fig. 2.1). They can interfere with the transport of essential nutrients within neurons. The presence and quantity of these tangles are linked to the extent of brain damage observed in individuals affected by AD [30].

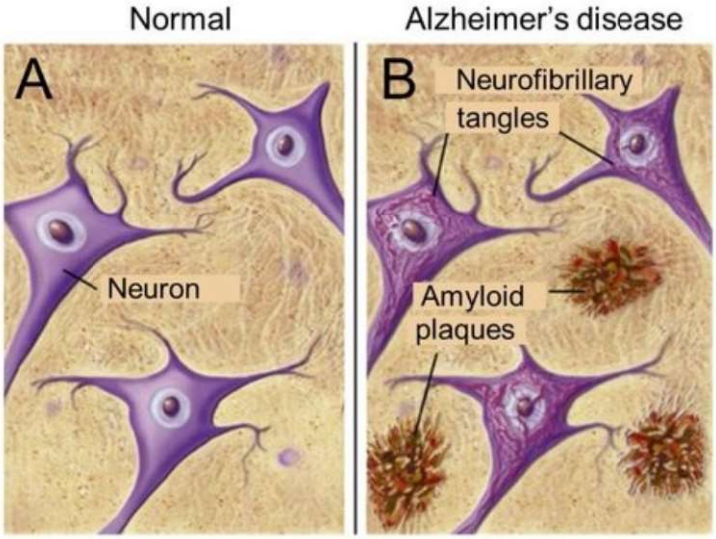

Figure 2.1: Major pathological hallmarks of AD are amyloid plaques and neurofibrillary tangles (B) that are absent in healthy brain tissue (A). (Sources: Google Images, BrightFocus Foundation, Alzheimer's disease: Amyloid Plaques and Neurofibrillary tangles).

According to this hypothesis, AD progression stems from the extensive accumulation of Aβ and neurofibrillary tangles in various brain regions. Initially, these accumulations manifest in specific areas such as the neocortex, hippocampus, and subcortical structures [146] before clinical symptoms arise. At this early stage, AD primarily affects the limbic regions linked to episodic memory, causing difficulties in retaining new information. Additionally, disruptions in connections between the frontal lobes and the hippocampus, coupled with hippocampal shrinkage, contribute to initial memory issues in AD patients [176]. As the disease advances, it encompasses other brain regions [28], impacting cognitive functions like executive functions, attention, visuospatial abilities, and language skills (Fig 2.2).

Diagnosing AD remains challenging due to its intricate nature and multifaceted

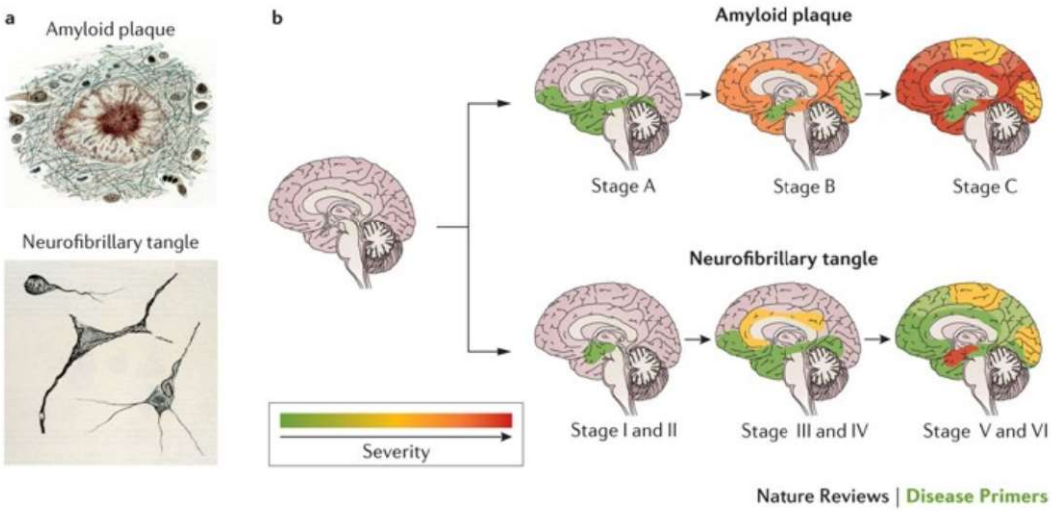

Figure 2.2: Evolution of AD in the brain: Plaques and tangles tend to spread through the cortex as AD progresses. (DOI: https://doi.org/10.1038/nrdp.2015.56)

causes. Age is a significant risk factor, alongside genetic elements, particularly evident in early-onset AD cases (occurring in individuals under 60) [207]. Reports indicate genetic associations in about 70% of AD cases [193]. Furthermore, vascular diseases, head injuries, infections, and environmental factors are linked to AD development. These various risk factors might collectively contribute to AD's pathology. Therefore, understanding their interactions and combined impact is crucial in comprehending the onset and progression of AD.

2.1.3 Progression of Alzheimer's Disease

AD is a chronic neurodegenerative disorders that can progress through five stages: preclinical AD, mild cognitive impairment due to AD, mild, moderate and severe dementia stages. In which:

- **Preclinical Stage**: This initial phase precedes observable symptoms [167], during which underlying brain changes linked to AD begin to emerge. It introduces the concept of Subjective Cognitive Impairment (SCI), where individuals notice a decline in cognitive abilities compared to their previous normal state[108], often years or even decades before visible symptoms appear.

- **Prodromal Stage**: Known as Mild Cognitive Impairment (MCI) due to AD, this phase exhibits measurable memory impairments, mild language difficulties, and concentration issues [63, 164]. Though functional abilities remain intact, these subtle changes are noticeable to close contacts. Studies suggest that within 5 to 10 years after an MCI diagnosis related to AD, approximately 30-50% of individuals progress to Alzheimer's dementia [134, 61, 7].

- **Mild AD Stage**: This phase showcases noticeable cognitive deficits, including memory and learning impairments. Memory troubles, disorientation, problem-solving difficulties, and psychological symptoms like compulsive behavior often surface. It's commonly when AD is diagnosed.
- **Moderate Stage**: Cognitive impairments become significantly pronounced, leading to heightened confusion and an increased need for assistance in daily tasks and personal care. Memory loss deepens, judgment weakens, and shifts in behavior and personality become apparent.
- **Severe Stage**: Marked by a severe decline in both cognitive and physical functions, individuals experience profound impairments in memory, thinking, and motor skills. Communication becomes severely limited, and independence in daily activities vanishes, making constant caregiver support necessary. On average, individuals diagnosed at this stage survive around 5 to 8 years post-diagnosis [89].

It is worth noticing that these stages offer a general framework, yet the symptoms experienced can differ significantly among individuals.

2.1.4 Alzheimer's Disease Diagnosis

Early and accurate diagnosis of AD is extremely importance as it offers several significant advantages. Primarily, it allows individuals and their families to prepare emotionally, financially, and practically for the challenges ahead. Understanding the disease in its initial stages aids in better disease management by enabling timely interventions, access to available treatments, and the implementation of strategies to slow down symptom progression. Moreover, comprehending the disease's progression and its early impact on daily life helps individuals and caregivers adapt, potentially enhancing the patient's quality of life and alleviating caregiver stress. Additionally, an early diagnosis allows participation in clinical trials for potential treatments, contributing to advancements in AD research and the development of new therapies.

However, achieving a reliable early diagnosis of AD remains challenging due to various factors. The pathological development of AD can begin up to two decades before noticeable clinical symptoms appear [61, 195]. Additionally, early symptoms are often mistaken for normal age-related changes, leading to delayed diagnosis[4]. Furthermore, AD progression varies significantly between individuals, resulting in a wide range of cognitive decline phases. As a consequence, diagnosing AD in the severe stages of the disease is typically easier than identifying it in its early stages.

Accurately diagnosing AD is challenging as definitive confirmation is only possible postmortem, through examination after death revealing specific brain damage related to the disease. In vivo diagnosis relies on a battery of clinical criteria, including neurological tests and medical recordings. The first official guidelines for diagnosing AD were established in 1984 by the National Institute of Neurological and Communicative Disorders and Stroke (NINDS) and the Alzheimer's Disease and Related Disorders

Association (ADRDA) [135], updated in 2011 by the National Institute on Aging and the Alzheimer's Association (NIAAA) [136, 197]. Currently, diagnosing the disease while the patient is alive depends on various clinical criteria, including neurological tests and medical recordings, yet definitive confirmation remains challenging.

Several specific tests, like the Mini Mental State Examination (MMSE)[141], the Montreal Cognitive Assessment (MoCA)[144], and other neurological tests like the Trail Making Test [10] and the clock-drawing test [192], have been proposed and used in clinical practice to assess cognitive function and diagnose AD. However, utilizing neuropsychological tests and medical recordings for AD diagnosis is time-consuming and requires experienced clinicians. An accurate and early diagnosis of AD holds immense importance, but diagnosing it in its early stages poses challenges when relying solely on clinical tests. To address this, AD biomarkers can be utilized.

According to [20, 203], biomarkers are evidence of any biological, pathogenic, or pharmacogenomic response when administered to any therapeutic change. In simple terms, a biomarker is a measurable feature that can be used to diagnose a physiological or pathological condition. For AD, an ideal biomarker would encompass being directed at the fundamental pathophysiology of the disease, serving as a marker of the disease's presence, exhibiting efficacy in prodromal and pre-clinical stages of AD, indicating the severity of the disease, reflecting treatment effectiveness, and being both inexpensive and non-invasive [113].

In the case of AD, different biomarkers have been described using diverse approaches [111, 133]. NIAAA recently grouped AD diagnostic biomarkers into three categories: beta-amyloid deposits (A), tau aggregates (T), and neurodegeneration (N) [105]. Assessment of the ATN profile is being done through biofluids (i.e., cerebrospinal fluid - CSF) or imaging (i.e., Positron Emission Tomography - PET). Efforts have also been deployed to explore simple and reliable markers for AD that would be less expensive than a PET scan and less invasive than a CSF measure. For example, biomarkers found in blood, saliva, urine; using eye tracking, gene data, etc. [82], can be considered as interesting biomarkers for further studies. It is important to know that the choice of a biomarker largely depends on cost and availability, and no single biomarker can, by itself, accurately diagnose a heterogeneous disorder such as AD [111]. A combination of several biomarkers yields better diagnostic results [133].

Neuroimaging serves as a type of biomarker that not only aids in better diagnosing brain disorders but also provides invaluable insights into the brains of live human subjects across various conditions. A key advantage of neuroimaging biomarkers is their higher level of spatiotemporal sensitivity compared to fluid biomarkers [133], offering an opportunity to stage the progression of diseases. Brain imaging modalities such as MRI and PET have become more common types of neuroimaging modalties for AD [111, 133, 82]. In the following sections, we will briefly describe several of the most common brain imaging biomarkers.

2.2 Introduction to NeuroImaging

2.2.1 (f)MRI and its use

2.2.1.1 Anatomical MRI

Magnetic resonance imaging (MRI) is based on the theory of nuclear magnetic resonance to generate structural images of inner organs or tissues with high quality. T1-weighted MRI is one of the most common types of MRI imaging modalities. It provides an anatomical view of the brain. T1-weighted MRI shows the brain anatomy with a high resolution and can clearly distinguish between gray matter and white matter. Therefore, tissue damage or loss can be estimated through structural MRI. By this T1-weighted MRI is an excellent method to assess brain atrophy.

Cerebral atrophy is considered a measure of neurodegeneration. It is correlated with tau deposition [222] and cognitive deficits [184]. In AD the progression of atrophy is well established, starting in the medial temporal lobe, then the temporal neocortex, associative parietal areas and frontal regions. On the other hand, visual and primary sensorimotor cortices remain relatively spared until late in the disease course [115]. Figure 2.3 shows the evident brain atrophy associated with AD compared to that of NC.

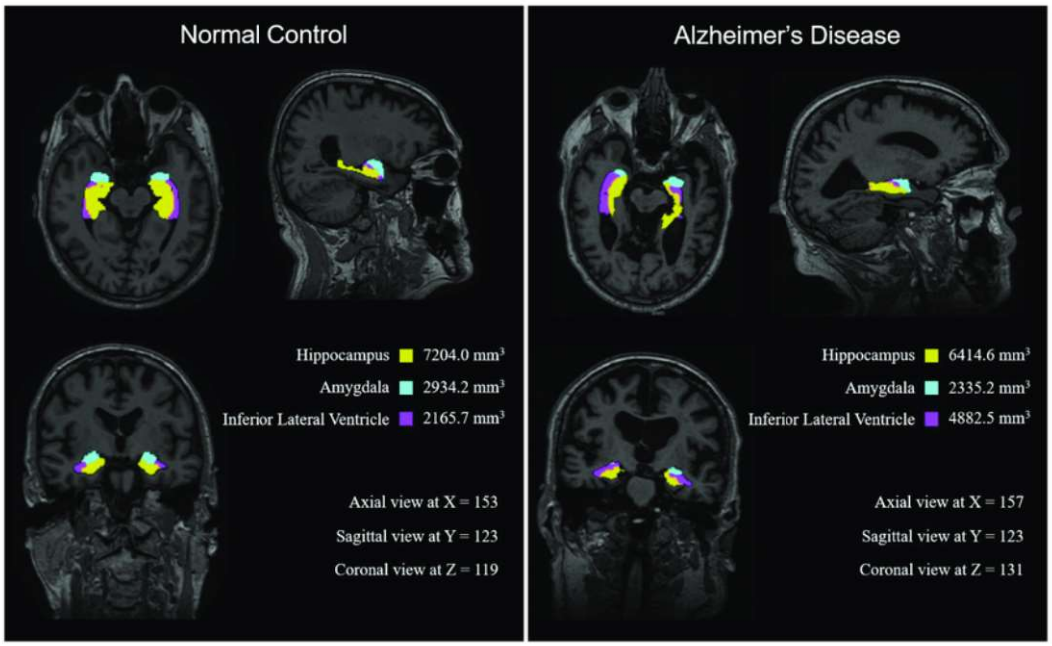

Figure 2.3: Two representative MRI images from the NC and AD groups (10.3390/brainsci11040453). Notably, the NC's hippocampus and amygdala are larger and the inferior lateral ventricle is smaller than that of the AD subject.

One advantage of using MRI is that it is widely available, non-invasive and relatively cheap compared to other modalities. Moreover, it is recommended to systematically

perform an MRI for evaluation of dementia or cognitive impairment , mainly for ruling out other possible causes, for example brain tumors. Therefore, in AD patients, MRI is acquired as part of the clinical routine examination.

2.2.1.2 Functional MRI

Functional MRI (fMRI) provides a measure of changes in blood oxygen levels that is an indicator of neuronal activity. fMRI is usually acquired during cognitive tasks or during a resting state (rs-fMRI). Functional connectivity between different brain regions can then analyzed. In particular, various studies based on rs-fMRI have evidenced alterations in the so-called "default-mode network" in MCI and AD patients [210, 79]. Moreover, the regions that showed connectivity alterations match quite well those with a heavier amyloid deposition in AD patients [196]. Figure 2.4 illustrates the lack of connectivity caused by AD.

Such techniques thus need to be evaluated across large, multicenter, datasets. fMRI is usually not part of a clinical routine MRI examination. It is also a technique very sensitive to subject movement and that may be difficult to harmonize across imaging centers.

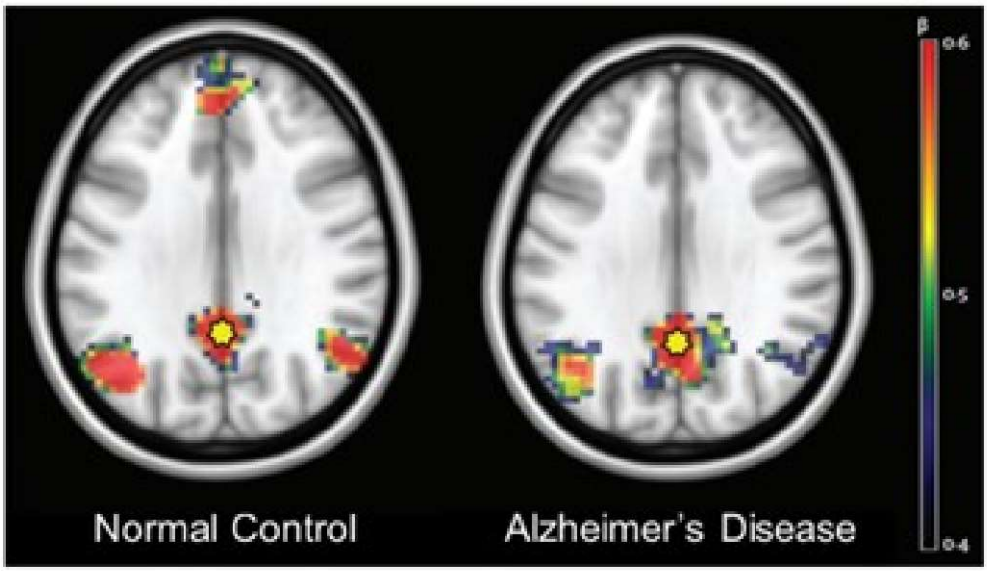

Figure 2.4: Differences of functional connectivity between NC and AD captured by rs-fMRI (axial view)

2.2.1.3 Diffusion MRI

Diffusion MRI (Diffusion tensor imaging - DTI) measures the diffusion of water along axons and can provide a representation of white matter fiber bundles in the brain. It is sensitive to the microstructural damage that may be present in the white matter bundles and offers different measures of its integrity. Then, it can be used to measure

the anatomical connectivity between brain regions. It allows detection of anatomical connectivity injuries that would not be detectable only with anatomical MRI.

Diffusion MRI allows distinguishing the progression pattern of the structural injury of white matter for AD (Fig. 2.5). Since the MCI stage, differences with control subjects are observed in the para hippocampal gyrus, temporal white matter, splenium of corpus callosum and posterior cingulum [44, 139, 96, 244]. Later, for AD, the damage is more severe in the previous regions and also includes the posterior regions [27, 139, 96, 145].

A drawback of diffusion MRI is that it is not usually part of a clinical routine MRI examination. Moreover, it provides high quality only on scanners with a magnetic field equal or above 3T and results in a longer acquisition time that can be troublesome for the patient.

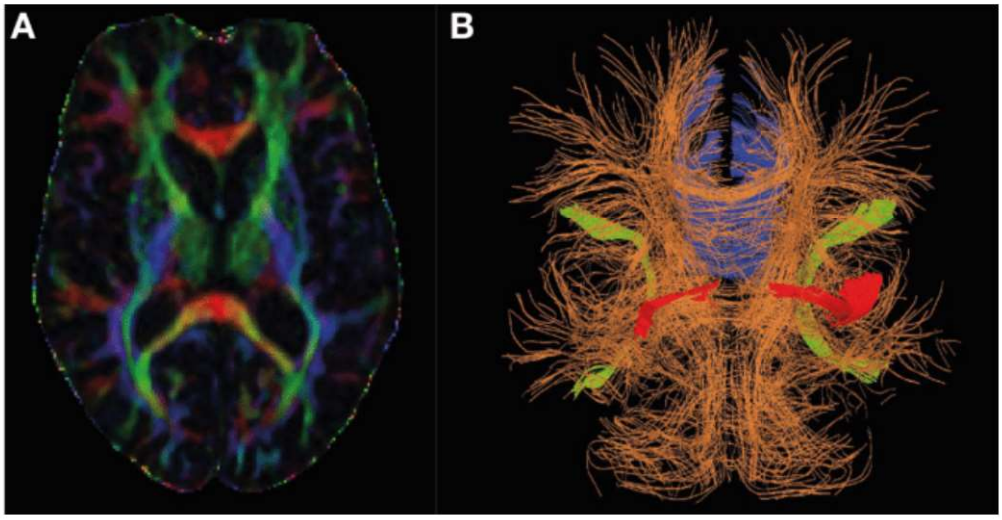

Figure 2.5: Diffusion tensor MRI in Alzheimer's research (DOI: 10.3389/fnagi.2016.00139). On the left is the directional information encoded by diffusion tensor imaging by color (red: left-right, green: anterior-posterior, blue: head-foot) and fractional anisotropy (high intensity implies high fractional anisotropy). On the right is an illustration of how the directional information may be represented in a deterministic tractography analysis, and shows the regions consistently identified and changing in AD (Blue: Corpus callosum, red: uncinate fasciculus, green: superior lateral fasciculus)

2.2.2 PET and its use

Positron emission tomography (PET) is a functional imaging technique that provides a representation of a given metabolic process through the detection of a positrone-mitting isotope that is bound to a biologically active molecule. Depending on the metabolic process in which this molecule is involved, a specific phenomenon will be observed. Although PET is not as widely available as MRI, is more expensive, and

involves injection of a radioactive tracer, it is the second most-widely used modality after anatomical MRI in AD-related ML studies. This is probably because PET provides information about the disease undetectable with MRI and because this modality is available in several publicly available research studies on AD such as ADNI.

2.2.2.1 FDG PET

Fluorodeoxyglucose Positron emission tomography (FDG PET) is an analog of glucose, the brain's main source of energy. A reduction of glucose metabolism indicates impairment of synaptic function. In AD patients, the hypometabolism is mainly found in the posterior cingulate gyri, precuneus, and parietotemporal association cortices [55, 15]. Regional hypometabolism is not a pathophysiological marker of AD, but considered a result of its degenerative processes. Fig. 2.6 shows an instance in which AD and MCI patients can be distinguished from NC by using FDG-PET scan.

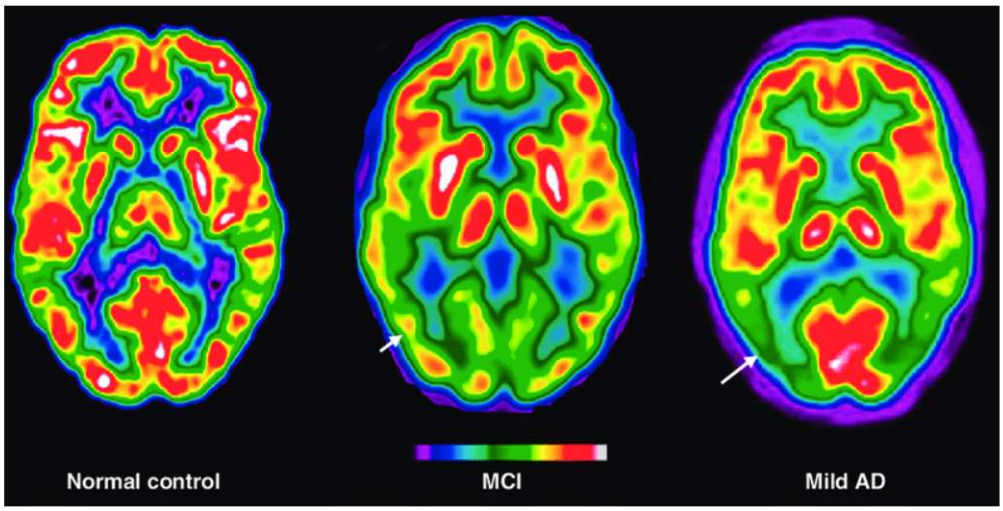

Figure 2.6: FDG-PET of control, MCI, and mild AD individuals (DOI: 10.2165/11595090-000000000-00000). FDG-PET show reduced uptake of regional CMRgl in the temporal-parietal cortex (shown as arrows) in MCI and AD patients. AD = Alzheimer disease; CMRgl = cerebral metabolic rate for glucose; FDG = [18F]-fluorodeoxyglucose; MCI = mild cognitive impairment; PET = positron emission tomography.

2.2.2.2 Amyloid PET

In amyloid PET, a marker that binds to the $A\beta$ protein is injected to the subject. This amyloid tracer shows the areas of the brain where there occurs an $A\beta$ deposition . The common tracers in use are the Pittsburgh Compound-B (PiB) [114], marked with 11C, and more recently the Florbetapir (AV-45) [43], Florbetaben [17], and Flutemetamol [219], marked with 18F.

The main advantage of the use of amyloid PET is detection of regions with $A\beta$ deposition, making it a pathophysiological marker of AD [62]. The deposition occurs very early in the disease course. It is both an advantage, as the pathological process can be detected decades before symptoms, and a drawback, as CN subjects with positive amyloid PET may not develop AD before a very long time or even not develop it at all. Fig. 2.7 illustrates differences between an NC subject and an AD patient in PIB-PET.

Amyloid PET in general is expensive to perform and it is not widely available, being currently mainly confined to research studies. Since the visual pattern of positivity is particularly striking, there have been few studies using ML with amyloid PET as a single modality to assist diagnosis [220]. More recently, PET tracers which can bind to tau protein deposits have been developed [223].

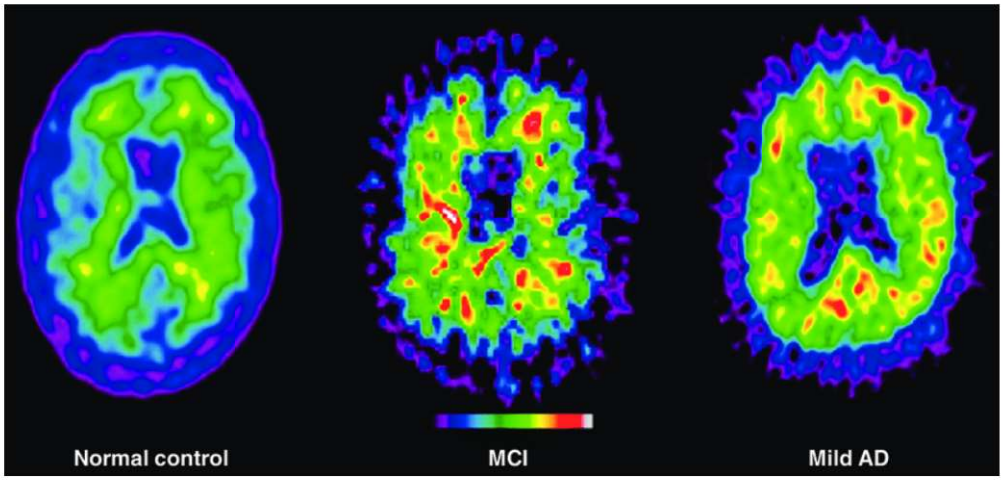

Figure 2.7: PIB-PET of control, MCI, and mild AD individuals (DOI: 10.2165/11595090-000000000-00000). The control shows negative PIB-PET with non-specific white matter retention but no cortical retention. The MCI image shows positive results with both white matter and cortical retention. The mild AD image shows typical positive PIB-PET with significant cortical retention in the prefrontal, posterior cingulate, and parieto-occipital cortices.

2.2.3 SPECT

Single photon emission computed tomography (SPECT) is an in vivo neuroimaging technique that uses gamma-emitting radiotracers to assess cerebral blood flow. The mainly used tracers include technetium-99m-labeled compounds, such as 99mTc-hexamethyl propylene amine oxime (99mTc-HMPAO) and 99mTc-ethyl cysteinate dimer (99mTc-ECD) [182]. These tracers are lipophilic and freely cross the blood-brain barrier in a manner proportional to the cerebral blood flow. A single photon is emitted in a blood-rich brain tissue, and then using tomography and image reconstruction

to form multiple azimuth sections and three-dimensional images. The changes in brain function are reflected by the measure of regional cerebral blood flow (rCBF). The majority of SPECT studies [39, 46, 106] have validated that evident deficits in perfusion can be observed in temporal and parietal regions in AD as compared to NC when using 99mTc-HMPAO and 99mTc-ECD as radiotracers. In addition, the reduction of rCBF in the posterior cingulate gyrus of AD patients conduces to the early diagnosis of AD, and can also be used to predict the conversion from MCI to AD [80]. Fig. 2.8 shows an instance in which AD patient can be distinguished from NC by using FDG-PET scan.

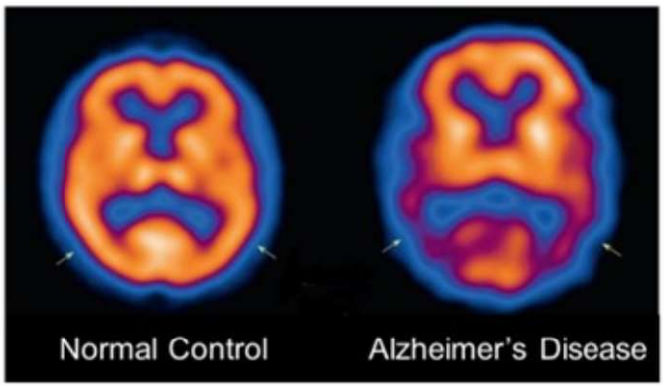

Figure 2.8: Differences between NC and AD captured by SPECT (axial view), from ADNI.

2.2.4 CT and its use

Computed tomography (CT) shows structures that are based on their ability to absorb the X-ray beam. Typically, bones, blood vessels and soft tissues will be clearly shown in a CT scan. CT uses a set of X-ray beams at various angles to produce cross-sectional images (aka tomographic slices) of specific areas of a scanned object. All captured images are then combined by computer-processed algorithms to create a complete volumetric (in 3D) representations of structures. CT is typically used to detect infarction, tumors, haemorrhage, and bone trauma. Due to ionizing radiation concern, CT of the head should be avoided for adults and delayed pending clinical observation in the emergency department for children. As a substition, magnetic resonance imaging (MRI) of the head provides superior information as compared to CT scans when seeking information about headache to confirm a diagnosis of neoplasm, vascular disease, posterior cranial fossa lesions, cervicomedullary lesions, or intracranial pressure disorders, yet it does not carry the risk of exposing patients to strong ionizing radiation of X-ray beams.

2.2.5 M/EEG and its use

Electroencephalography (EEG) shows brain activity under certain psychological states, such as alertness or drowsiness. Sensors which are called electrodes are placed on the skin around the head (aka scalp) to record electrical activity of the brain. EEG measures voltage fluctuations resulting from ionic current within the neurons. EEG is most often used to diagnose epilepsy which causes abnormalities in EEG readings [57]. It is also used to diagnose sleep disorders, depth of anesthesia, coma, encephalopathies, and brain death. EEG used to be a first-line method of diagnosis for tumors, stroke and other focal brain disorders [41], but this use has decreased with the advance of highresolution anatomical imaging techniques such as magnetic resonance imaging and computed tomography. Despite limited spatial resolution, EEG continues to be a valuable tool for research and diagnosis. It is one of the few mobile techniques that offers millisecondrange temporal resolution which is not possible with CT, PET or MRI.

2.2.6 Other modalities

Besides imaging, other biomarkers exist for the detection of AD. Cerebrospinal fluid provides several markers, the main ones being $amyloid\beta 1-42$ ($A\beta 1-42$), total tau protein , and phosphorylated tau protein [22]. These are pathophysiological markers that are specific of AD pathology. Nevertheless, they are invasive due to the lumbar puncture and can be relatively costly (the analysis is not expensive in itself but the patient is often hospitalized for a day to perform the lumbar puncture). Genetic factors that increase the likelihood of developing AD have been identified. The most important genetic factor is the allele e4 of the ApoE gene [132]. Besides ApoE4, many other genetic variants associated to AD have been discovered [107] but they have a weaker influence compared to ApoE4. Recently, blood-based markers have also shown promise to revolutionize the diagnostic and prognostic AD, as well as to improve the design of interventional trials [86, 119].

2.3 Machine Learning

2.3.1 Machine Learning pipeline

The process for applying machine learning to neuroimaging involves several critical stages, each playing a specific role in analyzing brain imaging data **[figure]**:

- Data Acquisition: This initial phase involves collecting data through various imaging techniques like MRI, fMRI, or PET scans.
- Data Preprocessing: Once the data is acquired, the preprocessing phase focuses on cleaning and enhancing it. This step includes removing noise, standardizing image sizes, and correcting any artifacts that might affect subsequent analyses.

- Feature Extraction: Following preprocessing, relevant features are extracted from the data. This involves isolating specific brain regions, extracting morphological or graph metrics, and representing them numerically for machine learning algorithms to process.
- Feature Selection: This phase is crucial for reducing dimensionality and improving model performance. It involves selecting the most relevant features to enhance the efficiency of the machine learning model while minimizing noise and overfitting.
- Training and Testing: Using labeled data, the model is trained and tested. Training involves feeding the algorithm with labeled examples to learn patterns and relationships within the data. Testing evaluates the model's performance on unseen data to assess its predictive capability.
- Decision Making: Finally, the trained model is used to make inferences or predictions based on new or unseen neuroimaging data. In clinical applications, this phase can aid in diagnoses, prognosis, or understanding brain-related phenomena based on the learned patterns. Continuous monitoring, refinement, and interpretability of the model results are vital to ensure its reliability and applicability in real-world scenarios.

2.3.2 Features

Neuroimaging data provides various types of features crucial for understanding brain structure and function in different ways.

- Voxel-based Features: These features are derived at the voxel level, representing the smallest identifiable elements in an image (volumetric pixels). Voxel-based analysis involves examining each voxel's intensity or characteristics within brain scans like MRI or fMRI. This method captures local information across the entire brain, allowing for detailed analysis of structural or functional differences in specific brain regions.
- Region-based Features: Region-based features involve segmenting the brain into anatomically defined regions of interest (ROIs). Statistical measures like mean intensity, volume, or shape features extracted from these regions provide insights into the characteristics of specific brain areas. These features are valuable for understanding localized variations or abnormalities within defined brain regions.
- Graph-based Features: Graph-based features explore the brain's connectivity patterns represented as graphs or networks. This approach focuses on relationships between brain regions, analyzing connections or pathways between nodes (regions) and their topological properties. Measures such as connectivity strength, centrality, or network efficiency derived from these graphs reveal

insights into brain connectivity and network organization, offering a broader understanding of brain function.

- Other Features: Neuroimaging also encompasses other feature types, including texture-based features capturing spatial patterns or arrangements within images, and time-series features derived from dynamic imaging modalities like functional MRI, representing temporal changes in brain activity. Additionally, advanced features may involve machine learning-derived representations like deep learning-based features, which automatically learn hierarchical representations from raw neuroimaging data, providing more abstract and high-level representations for analysis.

These diverse types of features extracted from neuroimaging data enable comprehensive analysis, offering insights into different aspects of brain structure, function, connectivity, and dynamics. Integrating multiple feature types often leads to a more holistic understanding of brain-related phenomena.

2.3.3 Dimensionality reduction

Feature reduction techniques are critical in handling high-dimensional data, aiming to decrease feature complexity while retaining essential information for model efficiency. Three primary approaches include filtering, wrapper-based, and embedding techniques, each with distinct advantages and drawbacks.

- Filtering Approach: These methods pre-process features based on statistical properties or relevance scores independently of the learning algorithm. They offer computational efficiency but may overlook feature interactions crucial for model learning and are less adaptive to specific algorithms.

- Wrapper-Based Approach: Wrapper methods integrate the learning algorithm into feature selection, evaluating subsets of features iteratively. They offer tailored selection based on model performance but can be computationally intensive and prone to overfitting, especially with large feature sets.

- Embedding-Based Approach: Embedding methods transform high-dimensional data into lower-dimensional representations while retaining critical information. They reduce dimensionality but may lose some interpretability and struggle with capturing complex nonlinear relationships.

Each approach has trade-offs in computational efficiency, interpretability, and model performance. Choosing the appropriate technique depends on dataset characteristics, computational resources, and specific task requirements, balancing computational complexity reduction with preserving essential information.

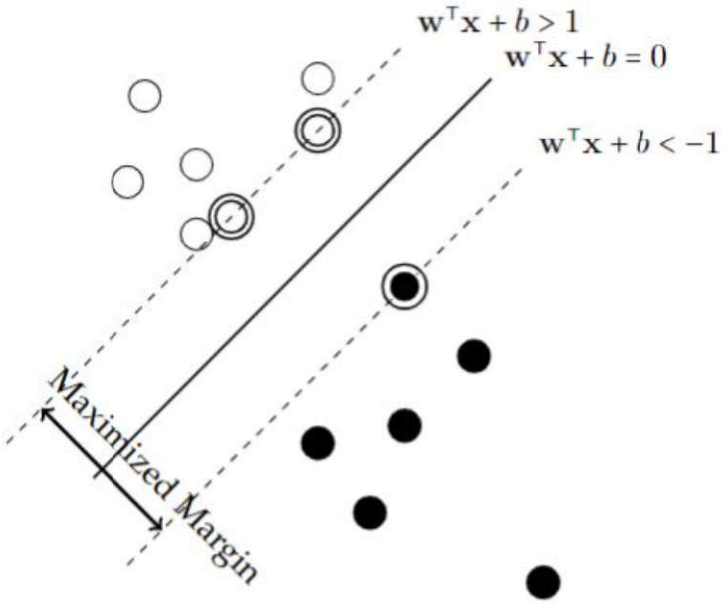

Figure 2.9: Support Vector Machine (SVM)

2.3.4 Learning approaches

2.3.4.1 Support Vector Machine

Support Vector Machines (SVM) is a supervised machine learning algorithm used for both classification and regression tasks. The primary objective of SVM is to find a hyperplane in an n-dimensional space (n being the number of features) that best separates different classes. In the context of classification, SVM aims to find the optimal decision boundary that maximizes the margin, which is the distance between the hyperplane and the nearest data points from each class, known as support vectors. SVM can handle linear and non-linear data through the use of different kernels, such as linear, polynomial, radial basis function (RBF), and sigmoid kernels, allowing it to efficiently classify data points into different classes by mapping them into higher-dimensional spaces. SVM is known for its ability to handle high-dimensional data and generalize well in complex scenarios, making it a widely used and versatile algorithm in various machine learning applications.

SVM is a widely adopted ML model that has found significant application in AD research. In particular, SVM is a supervised learning model that constructs a hyperplane or a set of hyperplanes in a high-dimensional or even infinite-dimensional space [226].

See Fig. 2.9 for an illustration.

2.3.4.2 Logistic Regression

Logistic regression is a widely used statistical method employed for binary classification tasks. Despite its name, it's actually a classification algorithm rather than a regression one. It's particularly suitable for scenarios where the outcome to be predicted belongs to one of two classes. This model applies a logistic function to predict the probability of an instance belonging to a particular class. It works by estimating the relationship between the independent variables and the binary dependent variable using a logistic function. By fitting a linear decision boundary, logistic regression effectively separates classes, making it a valuable tool in various fields such as finance, healthcare, and social sciences due to its simplicity, interpretability, and ability to handle linearly separable data.

2.3.4.3 Autoencoder

An autoencoder, despite being primarily used for unsupervised learning and dimensionality reduction, can be adapted for classification tasks. This neural network architecture consists of an encoder and a decoder, aiming to reconstruct input data. When used for classification, the autoencoder's encoder portion learns a compressed representation of the input features, reducing dimensionality. This compressed representation can then be utilized as input for a separate classifier, often a simple linear classifier or another neural network, to perform classification tasks. By leveraging the learned latent representations, the autoencoder aids in extracting meaningful features, effectively contributing to classification performance, especially in scenarios where the input data might be high-dimensional or noisy. However, this approach might require additional fine-tuning and might not always outperform traditional classifiers, yet it offers potential benefits in learning robust representations for classification.

2.3.5 Validation

In real tasks, there are usually a number of alternative learning models to choose among as well as several parameters to tune. We need to select the model and its corresponding parameter setting with the best performance. In order to achieve this goal, performance evaluation is necessary, which involves the commonly used and accepted measures and experiment designs.

2.3.5.1 Performance metrics

For a binary classification task, we can obtain different combinations of true labels and predicted labels, that are, true positive (TP) representing correctly identified positive labels, true negative (TN) representing correctly identified negative labels, in contrast to those incorrectly classified, false positive (FP) which implies negative labels incorrectly classified as positive ones and false negative (FN) implying positive labels incorrectly classified as negatives. Those concepts can be expressed intuitively

by a confusion matrix, as shown in Table 2.1 in which the summation of TP, TN, FP and FN is equal to the total number of samples.

		Predicted	
		Positive (P)	Negative (N)
Actual	Positive (P)	TP	FN
	Negative (N)	FP	TN

Table 2.1: Confusion Matrix for a Binary Classification

The mainly used measures for testing the performance of a binary classifier are computed based on the confusion matrix, including accuracy (ACC) or balanced accuracy (bACC), sensitivity (SEN), specificity (SPE) and area under curve (AUC) which is usually referred to a Receiver Operating Characteristic (ROC) curve. Accuracy is the proportion of samples that are correctly predicted, which is com- puted through,

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}$$

Sensitivity implies the proportion of correctly classified patients and it is expressed as,

$$SEN = \frac{TP}{TP + FN}$$

which is also known as true positive rate (TPR). Specificity means the proportion of control samples that are correctly classified, which is defined as,

$$SPE = \frac{TN}{TN + FP}$$

and $1 - SPE$ is referred to as false positive rate (FPR). However, for the case that the dataset is unbalanced, it would yield an optimistic evaluation if ACC is used as a measure to test the classifier's performance. Because the classifier would be biased and dominated by the majority data in that task. To address this problem, balanced accuracy is introduced and it is expressed as,

$$bACC = \frac{SEN + SPE}{2}$$

ROC curve reflects the relationship between the true positive rate (SEN) and the false positive rate ($1 - SPE$) with the change of the discrimination threshold of a binary classifier, as shown in Fig. 2.10. An ideal classifier would achieve 100% sensitivity and specificity, as a result, the upper left corner is with the best performance. Generally the metric AUC is applied to quantify the ROC and a higher value indicates better performance.

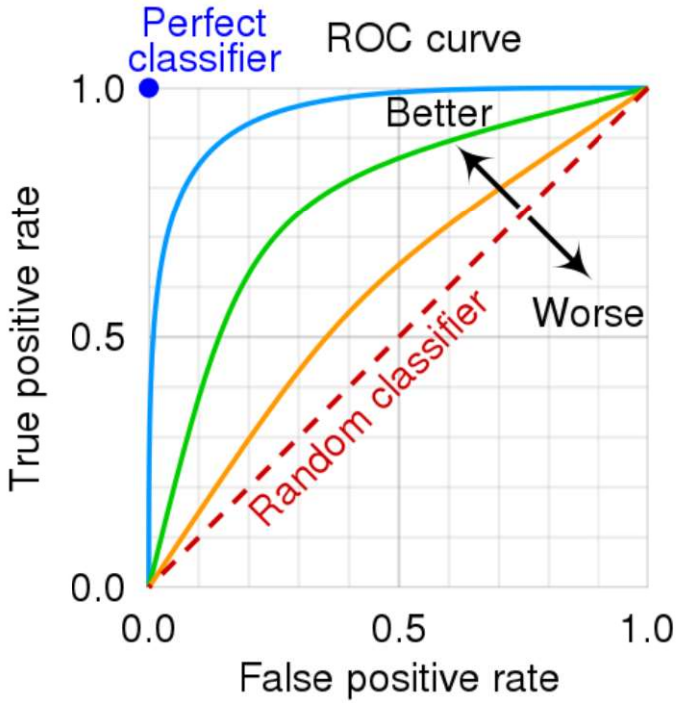

Figure 2.10: An instance of ROC curve (https://commons.wikimedia.org/wiki/File:Roc_curve.svg)

2.3.5.2 Cross-validation

A classifier is optimized based on the training dataset. An independent test set is therefore required to assess that classifier in terms of its application and generalization to new unknown data. If there is not enough labeled data available to create such a test set separately, a common method is to use the cross-validation (CV) technique [ref] in which the test set is also known as the validation set. There are roughly three types of CV methods, hold-out, k-fold, leave-one-out (LOO). Hold-out CV is easy to understand, which randomly divides the dataset into two groups, one is used for training, the other one is as the validation set. Note that the class percentage should be maintained for both training and validation sets when splitting the data, which is called stratification or stratified sampling. Obviously, the hold-out method only involves one round validation. It is normal to use such a method when the dataset contains a huge number of data since validation set could include a proper number of data to test the performance and meanwhile the computation consumption would be dramatically reduced. However, for the case that the dataset is not that large, which is common in medical data, hold-out method is not a suitable way to evaluate the classifier's performance because it is highly possible to achieve misleading results due to the small validation set. To solve this problem, k-fold method is commonly used.

K-fold CV takes advantage of different partitions of the dataset and performs multiple rounds of validation, in such way, a more accurate estimation of model prediction

performance will be achieved since it combines multiple results. In details, a dataset is usually partitioned with stratification into k equal-size subsets, then each subset could be the validation set and the others are the training set. Consequently, there are k combinations of training-validation set, and each of them is performed with the training-testing operation. After that, we can get k independent results, their average result is taken as the result of the cross-validation. Fig. 2.11 gives an insight into the procedure of cross-validation. To reduce the influence of randomness caused by data splitting, the k-fold CV is usually repeated multiple times, and the common configurations, including 10-times 10-fold CV and 5-times 2-fold CV.

LOO CV is a special case of k-fold CV in which k equals the number of samples in the original dataset and there is only one instance in each validation set. This method is usually applied when the dataset is particularly small, such as less than 50 samples.

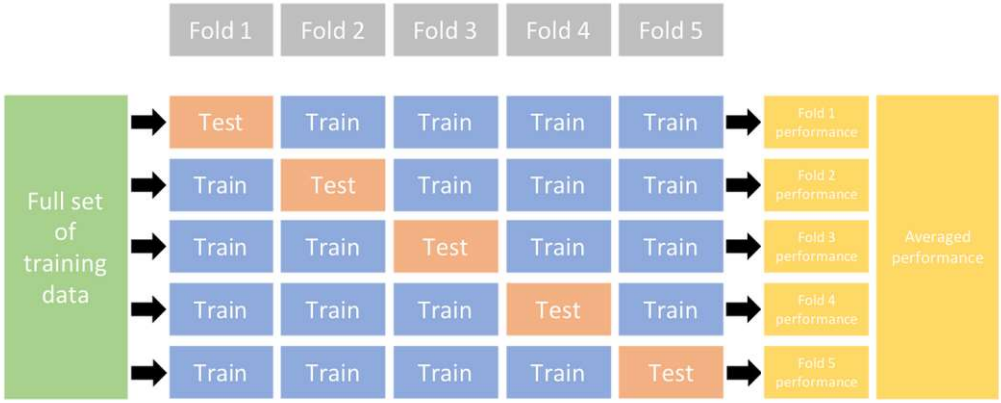

Figure 2.11: An illustration of 5-fold cross-validation.

2.4 ADNI dataset

2.4.1 Introduction

The Alzheimer's Disease Neuroimaging Initiative (ADNI) is a longitudinal multicenter study designed to develop clinical, imaging, genetic, and biochemical biomarkers for the early detection and tracking of Alzheimer's disease. ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. Four phases are included, which are ADNI 1 (from 2004 to 2010), ADNI GO (from 2009 to 2011), ADNI 2 (from 2011 to 2016) and ongoing ADNI3.

Different types of data are available on ADNI, including clinical data, genetics data, MRI, PET and biospecimens, as well as the corresponding longitudinal data. The data is free download for authorized investigators from Image and Data Archive (IDA) of Laboratory of Neuro Imaging (LONI) . In the following of this chapter, we mainly introduce the data acquisition from ADNI, data selection, and the processing procedure.

2.4.2 Data acquisition

FDG-PET images may be different due to different scanners, different shapes and positions of subjects' brains, etc. In order to make a consistent starting point for the subsequent data analysis, ADNI provided four types of pre-processed FDG-PET data, which are:

1. Co-registered Dynamic;
2. Co-registered, Averaged;
3. Co-reg, Avg, Standardized Image and Voxel Size;
4. Co-reg, Avg, Std Img and Vox Siz, Uniform Resolution.

In which:

- Type 1: Co-registered Dynamic Separate frames are extracted from the image file for registration purposes. Either six five-minute frames (ADNI1) or four fiveminute frames (ADNI GO/2) are acquired 30 to 60 minutes post-injection. Each extracted frame is co-registered to the first extracted frame of the raw image file (frame acquired at 30-35 min post-injection). The base frame image and the five co-registered frames (or all co-registered frames for the quantitative studies) are recombined into a co-registered dynamic image set. These image sets have the same image size (for example, $128 \times 128 \times 63$ voxels) and voxel dimensions (for example, $2.0 \times 2.0 \times 2.0 mm^3$) and remain in the same spatial orientation as the original PET image data. This is called native space.

- Type 2: Co-registered, Averaged This type of processed image set is generated simply by averaging the 6 five-minute frames (or the last 6 frames for the quantitative studies) of co-registered Dynamic image set. This creates a single 30 min PET image set still in native space. Type 1 and Type 2 data are only available for PET scans acquired under protocol 1 or 3 (ADNI 1 and ADNI 3).

- Type 3: Co-reg, Avg, Standardized Image and Voxel Size Each subject's Type 2 image from their baseline PET scan is then reoriented into a standard $160 \times 160 \times 96$ voxel image grid, having $1.5 \times 1.5 \times 1.5 mm^3$) voxel size. This image grid is oriented such that the anterior-posterior axis of the subject is parallel to ACPC line (AC: Anterior Commissure, PC: Posterior Commissure). This is referred to as AC-PC space. This standardized image then serves as a reference image for all PET scans on that subject. The individual frames from each PET scan (the baseline study as well as all subsequent studies (6-month scan, 12-month scan, etc.) are co-registered to this baseline reference image. By doing the co-registration from the original raw image data to a standardized space in a single step, only one interpolation of the image data is required, and thus resolution degradation by interpolation is kept to a minimum, and is the same for all scans. An averaged image is generated from the AC-PC co-registered frames and then

intensity normalized using a subject-specific mask so that the average of voxels within the mask is exactly one.

- Type 4: Co-reg, Avg, Std Img and Vox Siz, Uniform Resolution These images are the result of smoothing of the Type 3 images. Each image set is filtered with a scanner-specific filter function (can be a non-isotropic filter) to produce images of a uniform isotropic resolution of 8 mm full width at half maximum (FWHM), the approximate resolution of the lowest resolution scanners used in ADNI. Image sets from higher resolution scanners have been smoothed more than image sets from lower resolution scanners.

2.4.3 Data selection

Participants generally take several scans at different time point so as to track their situations. The first time of taking a scan is referred to as the baseline, then 6 months after the baseline, 12 months, 18 months, etc. As a result, longitudinal data is available in ADNI dataset. Considering the objective of this thesis is to diagnose AD and predict MCI conversion, only the baseline data is selected. But it is necessary to confirm that the selected data has not changed during its followup time. For example, we need to check whether the NC subject has converted to MCI or even AD within a certain month. If the NC subject has converted to MCI during the follow-up period, then the baseline scan from that subject will not be selected. Moreover, as many studies have done, MCI subjects are further classified into two different groups, progressive MCI (pMCI) and stable MCI (sMCI), in order to predict the conversion of an MCI subject. pMCI is the subject who progressed to AD in a certain period of time, while sMCI refers to the subject who keeps stable or reverses to NC. In fact, there is no uniform standard for setting the observation time. Some studies considered the data which does not change within 18 months, some considered data within 24 months or 36 months or even within the available scan time. Clearly, the longer the observation time is, the more challenging the prediction task.

In this thesis, we select the baseline FDG-PET data according to the following rules:

1. AD: subjects diagnosed as AD at the baseline and do not change within the follow-up time;
2. NC: subjects diagnosed as NC at the baseline and do not change within the follow-up time;
3. sMCI: subject diagnosed as MCI at the baseline and stay in the phase of MCI or revert to NC within the available scan time and the visit time is not less than 24 months;
4. pMCI: subjects diagnosed as MCI at the baseline and convert to AD and stay with AD in the available scan time.

Then according to the data selection rules and the diagnosis document, the experimental dataset which consists of 1048 subjects are obtained.

2.4.4 Data processing

After selecting the data, they are then further processed through the pipeline: reorientation (optional) → spatial normalization → intensity normalization → smoothing. These steps can ensure images in the same standardized space, then the subsequent analysis and comparison make sense. In this thesis, the MNI space is the standardized space for the experimental data.

All the procedures are implemented with SPM12 [ref], and the processing pipeline is shown in Fig. [ref].

The reorientation step in this thesis is to ensure the origin of the used image is roughly at AC point as the origin of MNI space is AC, which is an important step for the subsequent spatial normalization. If the image's origin has already pointed to AC, then this step is not needed. For the Type 3 image used in this thesis, the origin is at $(-160,-160,0)$, so the image is manually reoriented to its AC point, approximately at $(80,89.5,43.5)$ in its space. The spatial normalization step is to warp the image into MNI space and make the image have the same size with the template image. Here 'Old Normalize' unit and template image 'PET.nii' embedded in SPM12 are used. All the parameters $(-90-126-72)$ are defaults except the bonding box, which is reset as 9090108. Through this step, the FDG-PET image is with $2\times2\times2mm^3$ voxel size and $91\times109\times91$ matrix dimension. In addition, its origin is at $(46,64,37)$.

The intensity normalization is then performed through dividing each voxel intensity by the average global value. Thereafter, images are further smoothed by a Gaussian kernel with a full width at half maximum of 8 mm. Through the processing pipeline, 1048 baseline FDG-PET images constitute the experimental data, among which there are 237 subjects with AD, 242 subjects under NC and 569 subjects with MCI, including 209 pMCI and 360 sMCI. The demographic and clinical information of subjects is provided in Table 6.1, in which MMSE stands for the Mini-Mental State Examination.

Table 2.2: Demographic and Clinical information of subjects

Characteristic	CN	sMCI	pMCI	AD
Number of subjects	242	360	209	237
Female/Male	122/120	153/207	87/122	97/140
Age (Mean ± std.)	73.66 ± 5.66	71.73 ± 7.66	73.89 ± 6.88	75.00 ± 7.91
MMSE (Mean ± std.)	29.03 ± 1.20	28.20 ± 1.59	27.13 ± 1.71	23.19 ± 2.12

2.5 The state of the art

In the context of AD detection, SVM has been extensively employed, with a focus on the classification of AD using features derived from neuroimaging scans as input to the SVM classifier. Various adaptations of SVM have been applied for the problem of

Table 2.3: Performance of SVM Models for AD Detection

SVM Model	Data Type	ACC (mean ± std)	SEN (mean ± std)	SPE (mean ± std)
Linear SVM	MRI, PET, SPECT	90.11 ± 10.03	87.61 ± 10.56	89.29 ± 10.63
RBF SVM	MRI, PET	88.98 ± 7.28	90.11 ± 6.09	88.56 ± 7.37
Polynomial SVM	MRI	97.36 ± 3.60	93.53 ± 8.81	96.99 ± 2.67
Ensemble SVM	SPECT, MRI, PET	83.65 ± 16.84	79.54 ± 1.61	80.14 ± 14.3
Multikernel SVM	PET, MRI	84.91 ± 9.91	79.17 ± 19.62	83.64 ± 11.78
Lasso SVM	PET, MRI	83.68 ± 9.87	83.8 ± 15.55	81.48 ± 16.16
RF SVM	PET, MRI	88.79 ± 10.13	84.61 ± 8.41	83.75 ± 14.83

AD. Among these, the two most commonly used models are the linear SVM and the Radial Basis Function (RBF) SVM. The linear SVM employs a linear kernel, with the aim of identifying an optimal straight-line or hyperplane for effectively segregating data points within the original feature space. This approach proves highly effective when the underlying data distribution readily accommodates linear decision boundaries. Furthermore, its user-friendliness, robustness against overfitting, and minimal hyperparameter tuning requirements collectively position the linear SVM as one of the most efficient machine learning models for Alzheimer's disease detection. On the other hand, the RBF SVM employs a Gaussian radial basis function as the kernel. This choice is valuable for capturing intricate, non-linear relationships in the data by projecting it into a higher-dimensional space. Consequently, the RBF SVM has found extensive application in various studies focused on Alzheimer's disease detection.

Apart from the linear and RBF SVMs, several other SVM variations have proven to be valuable in addressing Alzheimer's disease (AD). For a summary of the performance of SVM models for Alzheimer's disease detection, please refer to Table 2.3.

Chapter 3

Autoencoder-based feature ranking

Summary

3.1 Introduction . . . 51
3.2 Autoencoder-based feature ranking for AD diagnosis . . . 52
3.2.1 Related work . . . 53
3.2.1.1 Feature selection . . . 53
3.2.1.2 Feature ranking & popular feature ranking methods . . . 53
3.2.1.3 Autoencoder-based feature selection . . . 54
3.2.2 Proposed method . . . 55
3.2.2.1 Dataset . . . 56
3.2.2.2 AutoEncoder . . . 57
3.2.2.3 Quantifying variable contributions in neural networks 58
3.2.2.4 Ranking methods . . . 59
3.2.3 Results and Discussion . . . 60
3.2.3.1 Results . . . 62
3.2.3.2 Discussion . . . 68
3.2.4 Conclusion . . . 74
3.3 Autoencoder-based feature ranking for MCI diagnosis . . . 75
3.3.1 Proposed Method . . . 76
3.3.1.1 Dataset . . . 76
3.3.1.2 2-phase feature ranking method . . . 77
3.3.1.3 Auto-encoder training . . . 77
3.3.1.4 Ranking methods . . . 78
3.3.1.5 Experimental Design and Evaluation Metrics . . . 78
3.3.2 Results and Discussion . . . 79
3.3.2.1 Results . . . 79
3.3.2.2 Discussion . . . 84
3.3.3 Conclusions and Future work . . . 87
3.4 Chapter Conclusion . . . 87

3.1 Introduction

As mentioned before, the ability to select the most important ROIs highly affected by AD is an extremely important task. From a medical standpoint, this capability aids in focusing on and tracking changes during various stages of AD, supporting doctors in better understanding and diagnosing the course map of the disease. From a machine learning perspective, training classifiers on the most informative features significantly improves classifier performance. Various techniques for feature selection have been proposed [ref1]. Generally, these techniques fall into three groups: filtering methods, wrapper-based methods, and embedding-based methods. Each has its own advantages and disadvantages. While filtering methods are simple, their results are often inadequate. Wrapper methods rely on a selected subset of features, making the division of features into subsets difficult and relying on rule-of-thumb-based approaches. Embedding approaches yield better performance, but they heavily depend on the models' capabilities to select the most relevant features.

Recently, there has been increasing interest in using deep learning models, particularly Autoencoders (AE), for selecting the most important features. AE, a type of neural network primarily used for feature extraction due to its powerful ability to learn effective features and its simplicity, is becoming more prevalent in machine learning and medical applications. Previous works introduced various AE architectures for selecting important features, showing interesting results but raising concerns about design complexity, optimal settings selection, and resource-intensive training.

In this work, instead of focusing on complex AE architectures for feature selection, we propose the concept of "single-but-multiple" AEs for feature selection. Specifically, we utilize multiple simple AE architectures with only hidden layers for feature ranking. Leveraging multiple simple AEs and proposing an algorithm to integrate results from these models, we expect that our method can help select the most important and stable features for AD.

The first section of this chapter introduces the AE-based feature ranking framework and its applications in AD classification tasks. In the second section, we propose improvements in training AEs, making our framework more efficient and powerful for selecting crucial features, applied to a more challenging task of MCI diagnosis. The last section compares the results and emphasizes the importance of brain regions in the course map of AD. Collectively, this chapter provides evidence for the effectiveness of AE in feature selection and the concept of "simple-but-multiple" models. The selected regions serve as a baseline for further comparisons and studies in subsequent chapters.

This chapter is based on our previous publications:

1. P. M. Tuan, T.-L. Phan, M. Adel, E. Guedj and N. L. Trung," AutoEncoder-based feature ranking for Alzheimer Disease classification using PET image," Machine Learning with Applications, vol. 6, p. 100184, 12 2021.

2. P. M. Tuan, N. L. Trung, M. Adel and E. Guedj, "AutoEncoder-based Feature Ranking for Predicting Mild Cognitive Impairment Conversion using FDG-PET

Images," 2023 IEEE Statistical Signal Processing Workshop (SSP), Hanoi, Vietnam, 2023, pp. 720-724, doi: 10.1109/SSP53291.2023.10208072.

3.2 Autoencoder-based feature ranking for AD diagnosis

One major challenge of PET-based classification is high dimensional dataset in which the dimensionality of image features is very high, while for getting high performance in classification tasks, a classifier should be trained on features that capture different patterns of structural degeneration [221]. To address this problem, dimensionality reduction could be used. Dimensionality reduction can be mainly categorized into two main groups: feature extraction and feature selection. The main difference between these groups lies in capability to retain the physical meanings of original features. While feature extraction creates a new feature set, feature selection selects a subset of features from the original features which means the physical meanings of features are maintained. Generally, feature selection may be done by selecting a subset of relevant features from the total number of features, or by ranking these features and selecting the most important ones. Although these two methods are related, they produce different results. While feature ranking provides information about the importance of individual features, the other one evaluates the importance of a subset of features, which means not all of features in the subset have high individual importance. Therefore, in the case of searching for the most effective brain region of interest (ROI) that is able to separate NC from AD, feature ranking would be more useful.

In this work, we present a method for ranking the effectiveness of ROI to separate NC from AD brains PET images. More specifically, inspired by the AutoEncoder (AE) neural network and derived from the basic idea that the ranking of features can be obtained based on the contribution of each feature to the neural network, we introduce a method for feature ranking based on the connection weights of multi-AutoEncoder models (AEs) that learned from the data.

In the literature, there are some researchers that use connection weights for quantifying the importance of features [77, 148, 54], however, one of the disadvantages of these studies is that the connection weights of neural networks are usually randomly generated. Random initialization often makes a significant difference between different experiments which leads to instability. To overcome this disadvantage, we consider multiple neural networks and use AE as a weight initialization method for neural networks.

Compared to previous studies, in this work:

- We propose a 2-phase feature ranking method based on AE and Connection Weights that addresses the issue related to randomly initialized weights.
- We evaluate the use of single simple AE architecture and multiple-view of them on the feature ranking task.

- Apply the proposed method to select the most effective ROIs for AD classification problems, the obtained result is competitive with other ranking methods.
- Moreover, while previous connection weights-based studies mainly worked with a small number of features (usually less than 20 features), here, we consider 90 features (90 ROIs).

The next section discusses the related work. Section 3 describes the dataset used in experiments and the proposed method. Section 4 presents experimental results and discussion. A summary and plans for future work are given in the last section.

3.2.1 Related work

3.2.1.1 Feature selection

Due to the ability to reduce the dimensionality of data and selecting the most informative subsets of features, feature selection can help to improve the performance of learning algorithms and reduce the demands for computational and memory resources. The literature on feature selection shows a variety of approaches. Depending on the availability of the labels, they can be categorized as supervised, unsupervised, and semi-supervised methods. While based on the selection strategies, they can be mainly divided into three major types [121]: Filter, Wrapper and Embedded methods. Filter methods rely on a ranking criterion to measure the importance of each feature. Wrapper methods usually select different subsets of features, measure their performance by a classifier to find the best subset, while embedded methods integrate feature selection into the learning process of model. Each one of these methods have their advantage and disadvantage. Filter methods are usually simple, fast, and computationally efficient; however, the selected features might be not good as wrapper's results. Wrapper methods have capability to select good subsets of features, however, they have been done in high computational complexity. Embedded methods have good the qualities of filter and wrapper methods, they are faster like filter methods while also considering the interaction with learning algorithm as wrappers methods. However, embedded methods are model dependent and only work with certain types of classifiers.

3.2.1.2 Feature ranking & popular feature ranking methods

In this work, our objectives were to find the most effective brain regions that could be used to separate NC from AD brains PET images, and to understand the importance of each individual region, therefore, we mainly focus on filter approach, working with feature ranking. A typical ranking process for feature selection consists of four steps:

1. Initialize set $\mathbf{F}$ to the whole set of p features. $\mathbf{S}$ is an empty set.
2. For each feature $f \in \mathbf{F}$ compute $J(f)$ coefficient.

3. Find feature f that maximizes $J(f)$ and move it to $\mathbf{S} \leftarrow \mathbf{S} \cup \{f\}, \mathbf{F} \in \mathbf{F} \{f\}$.

4. Repeat until the cardinal of $\mathbf{S}$ is p.

where $J(f)$ is a criterion function (different for any ranking algorithm) which gives a measure of dependency between feature (f) and classes (C).

According to the technical characteristics, feature ranking methods can be based on similarity score, statistics, information theory, or on some functions of classifier's outputs. Here, we briefly describe some popular and effective methods which will be used to compare with our proposed method. A more detailed and precise definition can be found in [121].

- **Fisher score** (similarity-based feature ranking) given class labels, this method tends to select features that have similar values within the same class and distinct values across different classes.

- **Conditional Mutual Information Maximization** (CMIM) (information theoretical based feature ranking) is a fast feature selection algorithm, it iteratively selects features while maximizing their mutual information between the feature being picked and class labels.

- **T-Score** (statistical based feature ranking) is used for binary classification problems. The basic idea of t-score is to assess whether the features make the means of two classes statistically different, which can be computed as the ratio between the mean difference and the variance of two classes.

- **Lasso Method** (sparse learning-based feature ranking) is a regression analysis method that performs both variable selection and regularization to enhance the prediction accuracy and interpretability of the statistical model it produces.

3.2.1.3 Autoencoder-based feature selection

In the last few years, some of autoencoder-based feature selection methods have been developed. In [85], authors proposed an unsupervised Autoencoder-based feature selection method, named AutoEncoder Feature Selector (AEFS). It incorporates autoencoder regression and $l_{2,1}$ regularization on the weights of the encoder to select a subset of most useful features. In [59], authors proposed an algorithm, called Agnostic feature selection (AgnoS), it combines an AE with structural regularizations to perform feature selection. More specifically, authors have applied $l_{2,1}$ regularization on the weights of the first layer, $l_{2,1}$ norm on the gradient of the encoder, l_1 norm on the slack variables of AE to delegate the feature selection task to AE. The recently proposed method, Concrete Autoencoder (CAE) [1], replaces the first hidden layer of AE with a "concrete selector" layer, which is the relaxation of a discrete distribution called concrete distribution, and features are selected based on their probability of connecting to the nodes of the concrete layer.

Compared to the above-mentioned methods, we use AE for feature selection in a different manner. Instead of embedding feature selection in the learning process of AE, we apply AE to initialize the weights of the networks. Specifically, we train multi-AEs with the number of hidden units varying. After that, we replace the decoder layer by softmax layer and tuning the network for classification task. The learned weights of an AE model can then be used to rank the input features. The final result is constructed from the ranked list of all multi-AEs. In this manner, we expect that the most important features selected by our method will be the ones which are both important to classifier and stable as the structure of the network changes.

3.2.2 Proposed method

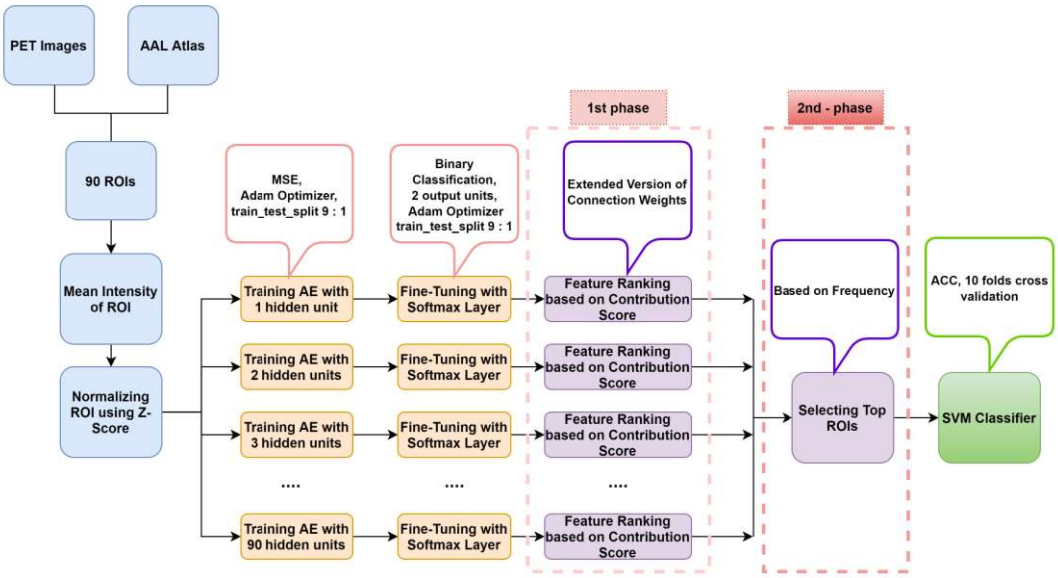

Figure 3.1: The flowchart of the proposed method.

Inspired by the idea of randomization approach for identifying variables that significantly contribute to network prediction in Olden's study [54, 147], in this paper, we use connection weights learned from multi-AutoEncoder models to rank ROIs according to the total contribution of ROIs to the networks. Here, instead of using initial random connection weights, we use AE as a way to initialize connection weights. For measuring the total contribution of each ROI, we use a modified version of Connection Weights algorithm. Based on the total contribution we proposed a method to rank and select the most important ROIs. In other words, we use the connection weights learned by AEs and a modified version of the Connection Weights algorithm to compute the total contribution of each ROI to the networks. After that, we use these contribution values as a criterion for ranking the ROIs. The flowchart of the proposed method is shown in Fig. 3.1. This flowchart can be divided into several blocks:

1. Data preprocessing (in blue).

2. Training AutoEncoder Models (in orange).
3. 2-phase feature ranking method (in violet).
4. Performance Evaluation (in green).

The different blocks appearing in this flowchart are detailed in the next sections.

3.2.2.1 Dataset

Data used in the experiment were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public–private partnership with the primary goal of testing whether serial MRI, PET, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of AD.

18F-FDG PET data was downloaded from ADNI and preprocessed before using for experiments. In ADNI dataset, each participant can take several scans at different time points, the first time of taking a scan is referred to as the baseline, then 6 months, 12 months, etc. Time points after the baseline are referred as follow-up time. Here, in this work, we mainly focus on AD diagnosis, therefore, only the baseline data is selected. Specifically, we select on data that meets the following criteria:

- AD: subjects diagnosed as AD at the baseline and do not change within the follow-up time.
- NC: subjects diagnosed as NC at the baseline and do not change over time.

In short, we only choose scans which are at baseline and correspond to subjects whose labels do not change for over at least 36-month period.

After downloading the data, images are spatially normalized to MNI template using SPM12 with $2 \times 2 \times 2mm^3$ voxel size, $91 \times 109 \times 91$ tensor dimension. The intensity normalization is then performed through dividing each voxel intensity by the average global value. After that, images are further smoothed by a Gaussian kernel with a full width at half maximum of 8 mm. All the procedures above are implemented with SPM12 [162]. Through the preprocessing pipeline, 479 baseline FDG-PET images constitute the experimental data, among which there are 237 subjects with AD, 242 under NC. The demographic and clinical information of subjects is provided in Table 4.1, in which MMSE stands for the Mini-Mental State Examination.

After that, each subject image is segmented into 90 ROIs using AAL atlas [214]. Each region is then characterized by mean value. To avoid biases, we use z-scores to normalize data:

$$z_i = \frac{x - \mu}{\sigma} \tag{3.1}$$

where μ, σ are the mean and standard deviation of the data.

Table 3.1: Demographic and clinical information of subject

Characteristic	**CN**	**AD**
Number of subjects	242	237
Female/Male	122/120	97/140
Age (Mean ± std.)	73.66 ± 5.66	75.00 ± 7.91
MMSE (Mean ± std.)	29.03 ± 1.20	23.19 ± 2.12

3.2.2.2 AutoEncoder

An autoencoder [91] is a specific form of an artificial neural network. The purpose of an autoencoder is to learn another representation of input data, in compressed for sparse representation. More specifically, an autoencoder is an unsupervised learning method that sets the target values to be equal to the input values.

Functionally, an autoencoder with a $h-dimension$ hidden layers contains two components: an encoder function $f(X) = \sigma_1(XW^{(1)})$ and a decoder that produces a reconstruction $\hat{X} = g(f(x)) = \sigma_2(f(X)W^{(2)})$, where σ_1, σ_2 are activation functions (which can be linear or non-linear ones such as sigmoid, ReLU, tanh, etc.) of the hidden layer and the output layer, respectively, $\Theta = \{W^{(1)}, W^{(2)}\}$ are weight parameters and denotes the parameter of the connection between the $i-th$ neuron in the $l-th$ layer and the $j-th$ neuron in the $(l+1)-th$ layer. The overall function of the autoencoder could be represented as $g(f(X))$. In the learning process, autoencoder is simply described as minimizing a loss function:

$$\mathscr{L} = \frac{1}{2m}||X - g(f(X))||_F^2 \tag{3.2}$$

where m is the number of samples and $||.||_F$ is the Frobenius norm of matrices. By optimizing the above function, we can obtain an autoencoder which compresses the observed matrix X as low-dimensional data $f(X)$ and outputs the decoded matrix $\hat{X} = g(f(X))$.

The more complex the network architecture and learning manner are, the higher possibility that the network overfits with data and the more challenging the assessment of feature contribution is, respectively. In our work, we consider simple AEs with 90 units in the input and output layer together with a varied unit number in the hidden layer (Fig. 3.2). By this way, we can evaluate the stability of the features with small changes in the network architecture before approaching more complex architectures. More specifically, we train 90 AE models with the number of hidden units varying from 1 to 90. For activation function, we select logistic sigmoid function: $f(x) = \frac{1}{1+e^{-x}}$ that likely relates to non-linear relationship among brain region connections; furthermore, in our test, sigmoid gives a better performance result. The connection weights learned from AEs are then used as initial weights for the classifier network.

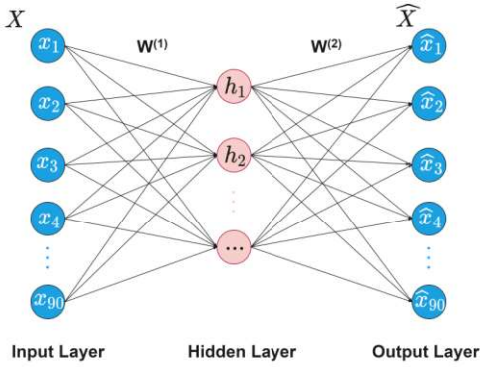

Figure 3.2: Simple Autoencoder structure.

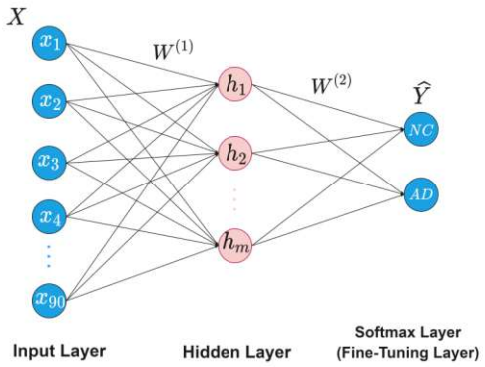

Figure 3.3: Fine-Tuning AutoEncoder.

3.2.2.3 Quantifying variable contributions in neural networks

To quantify variable contribution to neural the network, we fine-tune the learned AEs by removing the output layer of AEs, replacing it by the softmax layer. Here, we consider the case where the softmax layer has two units (Fig. 3.3). Then, we use a modified version of Connection Weights algorithm to calculate the contribution of each variable to the network.

Modified Version of Connection Weights algorithm: After fine-tuning AEs with softmax layer we get two matrices of weight parameters $W^{(1)}, W^{(2)}$, where $W^{(1)}_{ij}$ denotes the parameter of the connection between the $i-th$ neuron in the input layer and the $j-th$ neuron in the hidden layer, $W^{(2)}_{ij}$ denotes the parameter of the connection between the $i-th$ neuron in the hidden layer and the $j-th$ neuron in the output layer. The contribution that input unit i makes to output layer is:

$$\mathbf{c_i} = \sum_{j=1}^{p} |W^{(1)}_{ij}| \times \sum_{k=1}^{m} |W^{(2)}_{jk}| \tag{3.3}$$

where $\mathbf{c_i}$ is the contribution of the $i-th$ input unit made to the output layer; n, p, m are the number of units in the input layer, hidden layer, and output layer, respectively. Since the formula simply sums the products of two numbers, the computational complexity for measuring contribution of a ROI is $O(mp)$, and the computational complexity for all ROIs is $O(nmp)$. Note that the value of m is usually less than or equal to the number of labels of the data (in our case, m equals to 2), so it is not significant. Compared to Connection Weights algorithm, in the case that there is only one unit in the output layer, Eq. 3.3 is equivalent to:

$$\mathbf{c_i} = \sum_{j=1}^{p} |W^{(1)}_{ij}| \times |W^{(2)}_{j1}| \tag{3.4}$$

It is similar to Connections Weights algorithm, but instead of using the real value we use the absolute value of the connection weights. Therefore, our algorithm can be

viewed as a modified version of Connections Weights algorithm.

3.2.2.4 Ranking methods

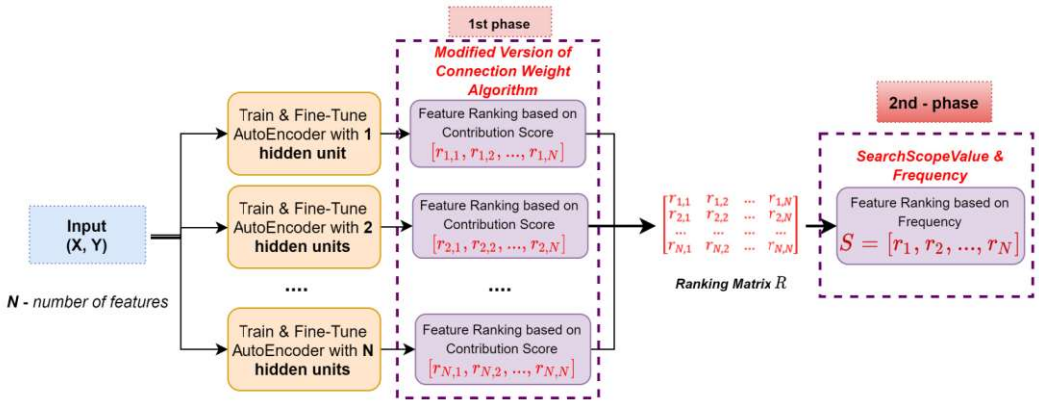

Figure 3.4: The flow chart of 2-phase feature ranking method.

In this section, we describe our proposed feature ranking method (Fig. 3.4), it can be divided into 2 phases:

First phase: Given N is the total number of features, we train and fine-tune N AEs whose number of hidden units varies from 1 to N. After that, we assess the contribution of each feature to each network and rank the features according to the descending value of contribution. More specifically, we will get a ranked list of N features for each autoencoder model. After these steps, we get N ranked lists. Stacking them together, we obtain a matrix with N rows, N columns, called ranking matrix **R**, in which each row corresponds to a ranked list of features which is given by a specific AE model. Element r_{ij} is in the $i-th$ row, the $j-th$ column of **R** corresponding to the ROI is ranked $j-th$ according to the $i-th$ autoencoder model. In our case, N equals to 90 ROIs, and after the first phase we obtain a ranking matrix **R** with 90 rows, 90 columns.

[**<empty citation>**] **Second phase:** Given a ranking matrix **R**, N is the total number of features, **S** is an empty set. We select the most important features and add them to **S** based on the following the steps (see Algrithm 1 for pseudocode):

1. Define $SearchScopeValue–k$: k can be any value in the range of $[1, N]$, however, we suggest selecting k is in the range from 20% to 50% of the total features.

2. Get the first k columns from **R** : With these k columns, we have a sub matrix of **R**, called $\mathbf{R}'$.

3. Determine the features appearing in $\mathbf{R}'$ and the corresponding number of occurrences, we obtain a list $f = \{(f_i, n_i)\}$ where $f_i \in \mathbf{R}'$, and n_i is the number of times f_i appears in $\mathbf{R}'$.

4. Sort f in descending order of n_i. For each $f_i \in f$ if $f_i \notin \mathbf{S}$, add f_i to the end of **S**.

5. If the cardinal of **S** is N or has selected enough features → Stop. In other cases, we continue to repeat steps 3-5 with the next k columns of **R**.

Algorithme 1 : 2nd ranking phase

Data : total number of features N, search scope value k, ranking matrix R with N rows and N cols

Result : Feature ranking list S

begin

$S \leftarrow \{\}$;
$start_col = 0$;
$end_col = k$;
while $n(S) \neq N$ **do**
Assign k columns from $start_col$ to end_col of matrix R to R';
Assign $f \leftarrow \{(f_i, n_i)\}$, where n_i is frequency of $f_i \in R'$;
Sort f by n_i;
for $f_i \in f$ **do**
if $f_i \notin S$ **then**
Append f_i to S;
$start_col = end_col$;
$end_col = end_col + k$;

3.2.3 Results and Discussion

We carry out experiments to compare the performance of our proposed method to other feature ranking methods on ADNI dataset. Comparative methods were selected based on the following criteria:

- Methods are based on ranking approach: To select most important features, we are interested in a ranking approach. It means we want to take into account the importance of all features, feature selection is then done by selecting the most important ones. It is important for medical applications as we want to know and evaluate the effect of each feature. In the literature there exists several studies related to autoencoder for feature selection, however, they have not been carried out for ranking purposes, therefore, we do not compare our proposed methods with these studies.
- Methods are well-known and popular.

Since Fisher, CMIM, T-Score and Lasso are well-known and representative for different ranking strategies such as: similarity-based, information theoretical based, statistical based, and sparse learning-based approaches [8, 121]; we decided to compare our proposed method with these methods.

For each ranking method, after ranking ROIs, the top-k ROIs are selected as candidate features for the classification. SVM is a popular and widely used technique for the detection and classification of patterns in medical images. In the literature, there are a lot of studies which use SVM as a classifier in AD diagnosis [36, 40, 75, 150, 151, 183]. Therefore, we use SVM as a classifier in this work. More specifically, we use 10-fold cross-validation for training and measuring the performance accuracy of the SVM classifier when the number of top-k of ROIs varies from 1 to 90. Given by the classifier, TP, TN, FP, and FN are true positive, true negative, false positive and true negative, respectively. The performance is measured based on ACCuracy, SPEcificity, SENsitivity and Area Under Curve (AUC) measures where:

$$\begin{aligned} ACC &= \frac{TP+TN}{TP+TN+FP+FN}, \\ SEN = \frac{TP}{TP+FN}, SPE &= \frac{TN}{TN+FP} \end{aligned} \tag{3.5}$$

AUC is the Area Under the Curve ROC (Receiver Operating Characteristics) determined by SEN and $1-SPE$. The higher value of these metrics indicates the better performance.

All experiments are implemented using scikit-learn [159] library. For experimental parameters, default values set by the libraries are used if not specified. Regarding dataset size, since we do not have much data, we will use same dataset source for both feature ranking and evaluating. More specifically, we divide the whole data into 2 sets (training/testing) in a 9:1 ratio and using these sets for training and evaluating performance of AEs. After that, we perform feature ranking steps. Finally, for evaluating the effective of ranking method, we use 10-fold cross validation with SVM classifier on whole dataset. Despite using the same source of data, feature ranking and evaluation are likely to be uncorrelated. While first process relies on the connection weights of AEs and the stability of features when the structure of AEs changes slightly, the latter considers the subset of features on SVM classifier.

3.2.3.1 Results

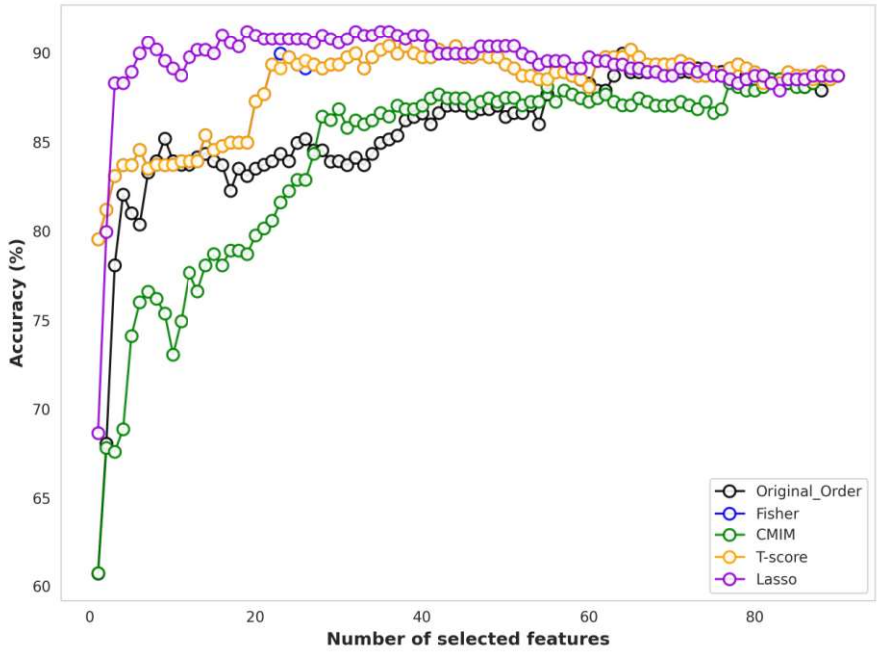

Figure 3.5: Performance of 4 baseline methods with different numbers of ROIs.

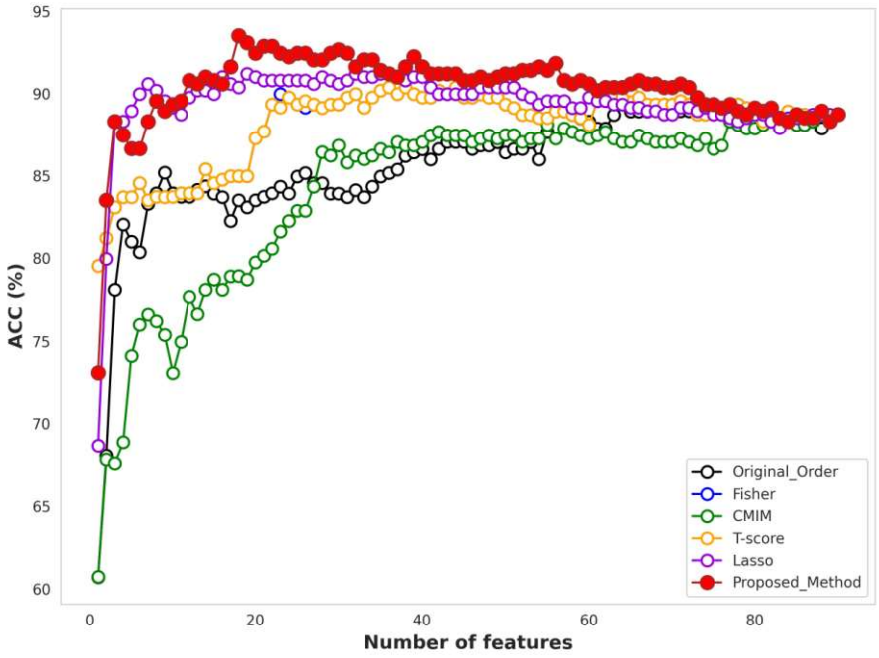

Figure 3.6: Comparing performance between proposed method and the other methods under different numbers of ROIs.

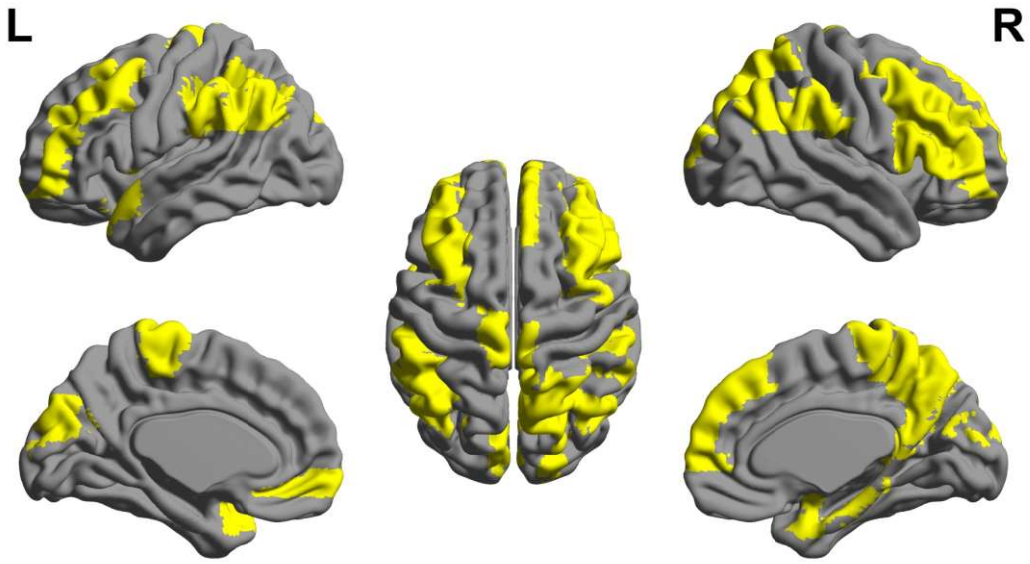

Figure 3.7: The top-20 ROIs selected by proposed method on 3D brain image.

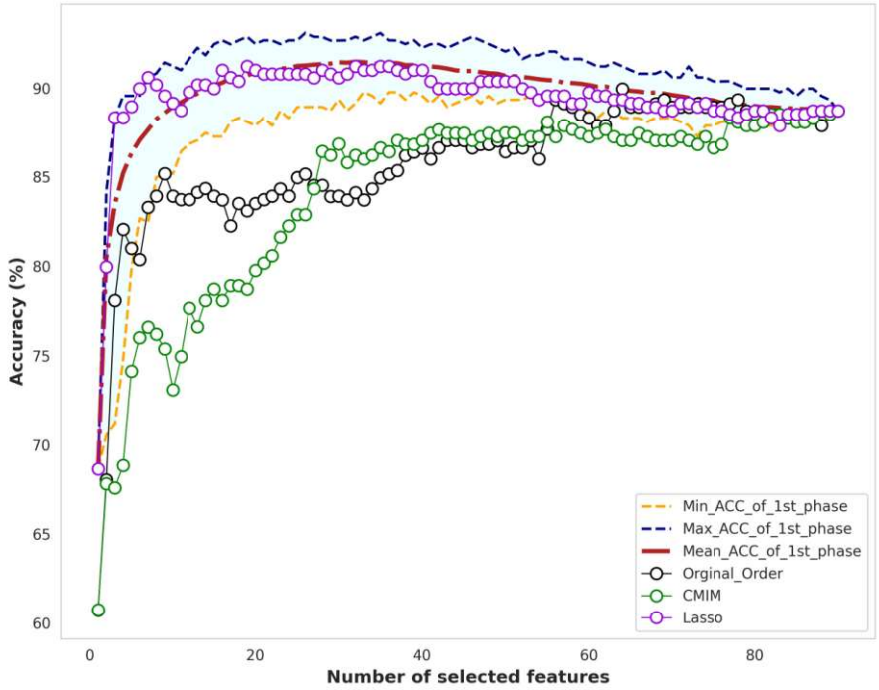

Figure 3.8: Comparing performance between 1st phase of proposed method and the other methods under different numbers of ROIs.

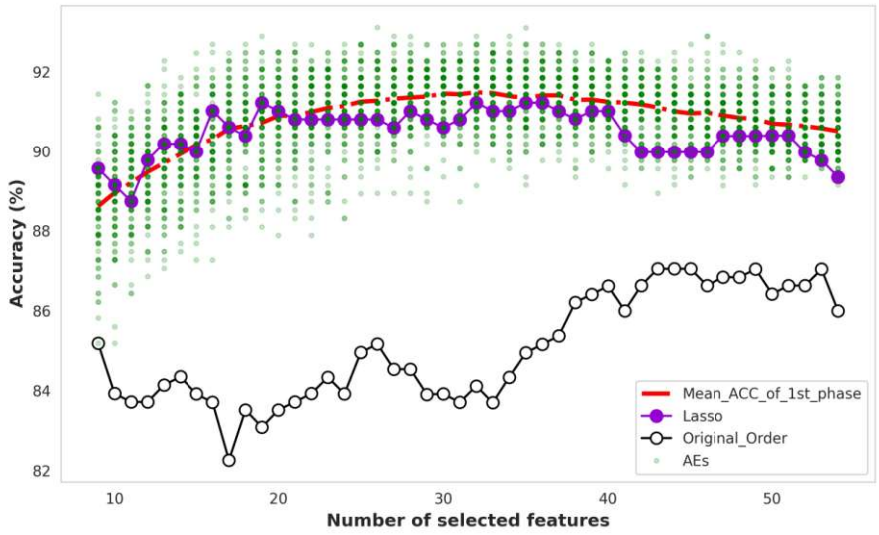

Figure 3.9: Comparing performance between 1st phase of proposed and Lasso method when the number of selected features is in range of [9,54].

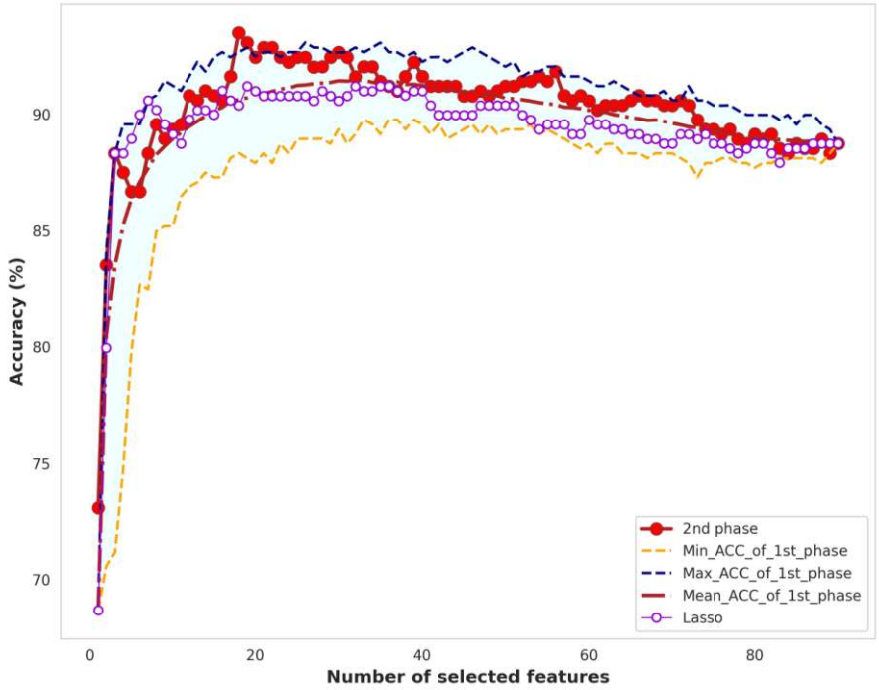

Figure 3.10: Comparing performance between the first and the second phase of proposed method under different numbers of ROIs.

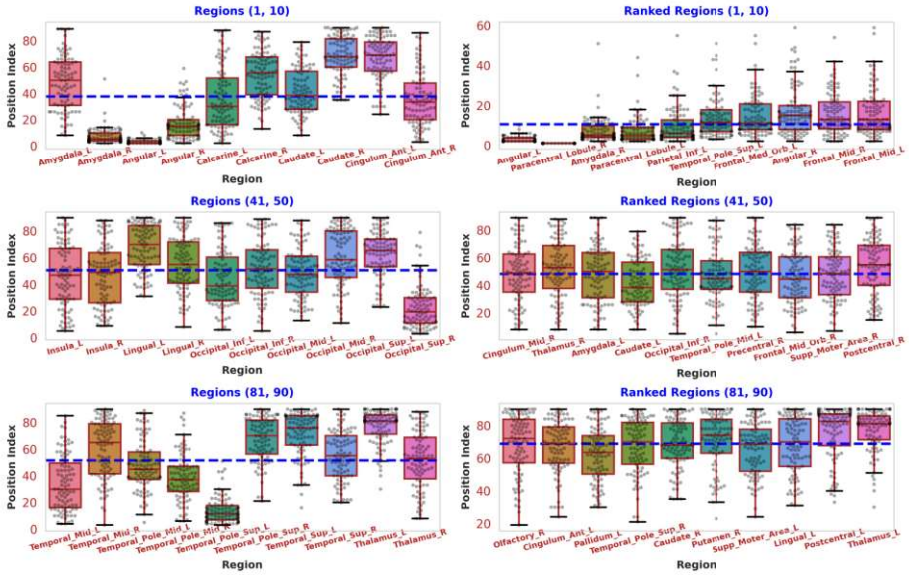

Figure 3.11: Comparing position Index of the all ROIs before and after the 2nd phase.

Firstly, we compare the accuracy of the SVM classifier between the original ordering of features and the ranked orders according to the selected ranking methods (see Fig. 3.5 and Table 3.2). We can observe that feature ranking methods significantly improve the average accuracy. From 1 to 20 features, in the best cases, Lasso has the most dramatically increased ACC and outperforms the original by 7.5% of ACC and by a similar amount of SPE and SEN. Meanwhile, Fisher and T-test perform similarly to, and CMIM is lower than the original. In between 20 and 40 features, Lasso has a stable ACC while others continue to reach their maximum ACCs. To all methods, we observe a gradual decline in ACC starting from 40 features, which means more features added from this point not only do not improve the classification performance, but also degrade the former good feature space.

Moreover, Table 3.3 shows the overlap rate between 4 baseline methods. We can observe that Lasso shares around half of top-k ROIs (40%–50%) in common with Fisher and T-test at almost all cases, while CMIM slowly keeps up with them and reaches the adequate mutual feature sets at top-40 ROIs. Based on the classification performance and the similarity of top-k selected features, Lasso can be considered as the best one among these methods. Therefore, we use Lasso as the main comparing method for the next experiments

Table 3.4 presents 10 shared ROIs when selecting the $top-k$ of ROIs based on Fisher score, T-test and Lasso ranking methods.

Next, we compare our proposed method with the methods above. Here, we choose the value of equals 18 (20% of the total ROIs) and conduct experiments. Fig. 3.6 and Table 3.5 compare the accuracy of the SVM classifier between our proposed method and the other methods. We can observe that the proposed method significantly

improves the accuracy of SVM classifier when compared to original order and other ranking methods. Moreover, in the case of selecting top-18 of ROIs (approximately 20% of the total features), the proposed method gives the highest accuracy (93.53%), and even outperforming Lasso (by 3.14%). Not only accuracy, from Table 3.2, Table 3.6, we can observe that other measurements are also improved, especially the case of SPE. Note that, from a medical aspect, the rate of true negative detection is also extremely important to avoid unnecessary treatment to healthy control group.

Table 3.7 compares the overlap rate between the proposed method with the other feature ranking methods. We can observe that the overlap rate between our method and CMIM is quite low (only 5% when selecting 20 ROIs) while its value when compared to Lasso is always greater than 50%, it means that our proposed method is capable of picking out the most important ROIs like the one obtained by Lasso. Moreover, the average accuracy of our method in most cases is always close to or higher than the results obtained by Lasso which means that it also has the ability to select the other effective ROIs that Lasso could not do.

Table 3.8 presents the top-20 of regions selected by the proposed method, these ROIs can be viewed as potential FDG-PET indicators for AD/NC classification tasks. The bolded ROIs correspond to the ROIs that also appear in the Table 3.4. Looking at the positions of these top-20 ROIs on AAL, we can observe that they mainly belong to Frontal and Temporal lobes (Fig. 3.7). This obtained result seems to be consistent with FDG-PET findings regarding the role of temporoparietal regions and frontal cortex in AD [26].

Table 3.2: Performance of 4 baseline ranking methods

Top-k ROIs	Original				Fisher				CMIM				T-test				Lasso			
	ACC	SEN	SPE	AUC	ACC	SEN	SPE	AUC	ACC	SEN	SPE	AUC	ACC	SEN	SPE	AUC	ACC	SEN	SPE	AUC
1	60.7	41.6	79.1	61.3	79.5	72.2	87	84.5	60.8	47.4	74.3	65.4	79.5	72.2	87	84.5	68.7	54.3	82.7	71.9
2	68.1	59.4	76.5	76	81.2	73.2	88.9	86.8	67.8	62	73.5	72.9	81.2	73.2	88.9	86.8	80	75.6	84.2	86.9
5	81	74.8	86.8	91	83.7	81.8	85.6	90.8	74.1	73.8	73.9	79.1	83.7	81.8	85.6	90.8	89	86.9	90.7	95.5
10	83.9	81.7	86.2	91.9	83.7	80.6	86.9	92.3	73.1	71.3	75	83.5	83.7	80.6	86.9	92.3	89.2	87.4	90.3	96.4
18	83.5	82.3	84.8	91.5	85	82.9	86.6	92.4	78.9	73	84.8	88.1	85	82.9	86.6	92.4	90.4	89.5	90.7	96.7
20	83.5	81.7	85.2	91.4	87.3	84.8	89.4	94.7	79.8	75.1	84.2	88.6	87.3	84.8	89.4	94.7	91	90.2	91.2	96.3
27	84.5	82.4	86.7	92.5	89.4	85.9	92	95.5	84.3	80.2	88.6	91	89.4	85.9	92	95.5	90.6	89.3	91.1	96.1
36	85.2	83.3	87.1	93	90.4	89	91	95.6	86.4	84.7	87.9	93.8	90.4	89	91	95.6	91.2	90.2	91.6	96
45	87.1	85.7	88.3	93.6	89.8	88.7	90.2	95.5	87.5	86.4	88.3	93.9	89.8	88.7	90.2	95.5	90	89.4	89.9	95.7
60	88.3	87	89	95	88.1	86.1	90	95.6	87.3	85.4	89	94.8	88.1	86.1	90	95.6	89.8	88	91.1	95.5
80	88.5	86.7	89.8	95.4	88.9	87.7	89.8	95.6	87.9	86.1	89.5	95.6	88.9	87.7	89.8	95.6	88.7	86.8	90.3	95.5

3.2.3.2 Discussion

In this paper, we present a 2-phase feature ranking method. The main ideas of this approach rely on quantifying the contribution of variables to multi-AutoEncoder models. In this section, we will verify the effectiveness of each phase in the proposed method.

After the first phase, we obtain a ranking matrix $\mathbf{R}$ of m rows, n columns, where n is the total number of features and m is the total number of AE trained models (here, in our case, $m = n = 90$). The $i-th$ row corresponds to a ranking list r_i of n features,

Table 3.3: Overlap rate when comparing 4 baseline ranking methods

top-5

	Fisher	**CMIM**	**T-test**	**Lasso**
Fisher	1			
CMIM	0	1		
T-test	1	0	1	
Lasso	0.4	0	0.4	1

top-10

	Fisher	**CMIM**	**T-test**	**Lasso**
Fisher	1			
CMIM	0	1		
T-test	1	0	1	
Lasso	0.3	0.2	0.3	1

top-20

	Fisher	**CMIM**	**T-test**	**Lasso**
Fisher	1			
CMIM	0	1		
T-test	1	0	1	
Lasso	0.5	0.2	0.5	1

top-30

	Fisher	**CMIM**	**T-test**	**Lasso**
Fisher	1			
CMIM	0.2	1		
T-test	1	0.2	1	
Lasso	0.53	0.33	0.53	1

top-40

	Fisher	**CMIM**	**T-test**	**Lasso**
Fisher	1			
CMIM	0.3	1		
T-test	1	0.3	1	
Lasso	0.45	0.48	0.45	1

top-50

	Fisher	**CMIM**	**T-test**	**Lasso**
Fisher	1			
CMIM	0.5	1		
T-test	1	0.5	1	
Lasso	0.52	0.58	0.52	1

Table 3.4: Shared rois between lasso, fisher, t-test when selecting top-20 rois

	Top ROIs
1	Angular_R
2	Angular_L
3	Cingulum_Post_L
4	ParaHippocampal_R
5	Parietal_Inf_R
6	Postcentral_R
7	Rolandic_Oper_R
8	Parietal_Inf_L
9	Precuneus_L
10	Paracentral_Lobule_R

Table 3.5: Comparing accuracy of SVM classifiers between proposed method with baseline methods

Top-k ROIs	**Accuracy of SVM Classifier**					
	Proposed Method	**Original Order**	**Fisher**	**CMIM**	**T-test**	**Lasso**
1	73.08	60.73	**79.54**	60.75	**79.54**	68.67
2	**83.52**	68.06	81.21	67.84	81.21	79.96
5	86.65	81	83.71	74.11	83.71	**88.95**
10	**89.36**	83.93	83.72	73.07	83.72	89.15
18	**93.53**	83.52	84.97	78.92	84.97	90.39
20	**92.48**	83.52	87.28	79.75	87.28	91.02
27	**92.06**	84.54	89.36	84.34	89.36	90.6
36	**91.23**	85.16	90.4	86.43	90.4	91.23
45	**90.81**	87.05	89.77	87.46	89.77	89.98
60	**90.6**	88.3	88.1	87.27	88.1	89.77
80	**89.14**	88.52	88.93	87.89	88.93	88.73
90	**88.73**	**88.73**	**88.73**	**88.73**	**88.73**	**88.73**

Table 3.6: Detailed performance of SVM classifiers for the proposed method

Top N ROIs	**Performance of SVM Classifier**			
	ACC	**SEN**	**SPE**	**AUC**
1	73.08	56.05	89.39	78.22
2	83.52	78.66	87.87	89.17
5	86.65	83.1	89.95	93.4
10	89.36	86.2	92.06	96.18
18	**93.53**	**91.93**	94.45	97.27
20	92.48	89.74	94.53	97.12
27	92.06	88.99	**94.63**	**97.33**
36	91.23	88.13	93.62	97.03
45	90.81	88.23	92.99	96.71
60	90.6	89.26	91.68	96.32
80	89.14	86.79	91.23	95.66
90	88.73	87.33	89.84	95.38

Table 3.7: Overlap rate when comparing proposed method with baseline methods

Top N ROIs	**Performance of SVM Classifier**			
	Fisher	**CMIM**	**T-test**	**Lasso**
1	0	0	0	0
2	0.5	0	0.5	0.5
5	0.2	0.2	0.2	0.6
10	0.2	0.1	0.2	0.5
18	0.28	0.06	0.28	0.56
20	0.3	0.05	0.3	0.5
27	0.37	0.26	0.37	0.52
30	0.37	0.27	0.37	0.5
36	0.47	0.39	0.47	0.56
40	0.55	0.42	0.55	0.55
45	0.56	0.49	0.56	0.58

Table 3.8: Top 20 rois selected by proposed method

	Top ROIs
1	**Angular_L**
2	**Paracentral_Lobule_R**
3	Amygdala_R
4	Paracentral_Lobule_L
5	**Parietal_Inf_L**
6	Temporal_Pole_Sup_L
7	Frontal_Med_Orb_L
8	**Angular_R**
9	Frontal_Mid_R
10	Frontal_Mid_L
11	Parietal_Sup_R
12	SupraMarginal_L
13	Frontal_Inf_Oper_R
14	**ParaHippocampal_R**
15	Frontal_Sup_Medial_R
16	Occipital_Sup_R
17	Cuneus_L
18	Precuneus_R
19	Frontal_Inf_Tri_R
20	SupraMarginal_R

which is given by assessing the contribution of each feature to the $i-th$ AE model. In order to evaluate the effectiveness of first phase, we consider the effectiveness of each AE model for ranking task. Given the number of selected ROIs – k varying from 1 to 90, with each value of k, we select top-k subset of features corresponding each list, using them as candidate features for training m SVM classifiers and comparing their accuracy together. We then compute the minimum, maximum and the average accuracy values of these classifiers. The obtained result is shown in Fig. 3.8, the dashed orange, blue and red curves represent the minimum, maximum and average accuracy value, respectively. It is easy to observe that the accuracy in most cases is always higher when compared to the results obtained by unranked list and CMIM method. The average accuracy is always close to or higher than the results obtained by Lasso method. It means the first phase of our proposed method can significantly help improve the accuracy of classifiers.

At a closer look in the range of [9,54] (10% to 60% number of ROIs), as shown in Fig. 3.9, the difference between minimum and maximum achievable accuracy points out the variation of AEs' capability for ranking task. Based on this, it can be seen that each AE has its own ability for selecting the most important features. In addition, the accuracies of all SVM classifiers are very close to ones of Lasso method and higher than ones of the original order. It means that each AE has a good capability for ranking task, and in most cases, it is possible to use a single AE while still guaranteeing that the obtained result can help to improve the accuracy of classifiers. This observation could lead to a possibility of multiple-view combination for an overall performance improvement of the proposed method, which will be confirmed in the 2nd phase.

Next, we will consider the effectiveness of the second phase. Given the number of selected ROIs varying from 1 to 90, we use ranking list from the 2nd phase for training SVM classifiers. The obtained result is shown in Fig. 3.10; the red curve depicts the 10-fold cross-validation accuracy of classifiers when the number of selected ROIs changes. It can be seen that the second phase of the proposed method significantly improves accuracy when compared to the average results obtained by the first phase and by Lasso method. In most cases, the obtained results are always higher than those obtained by Lasso method. When selecting top-18 ROIs, its value is even higher than the maximum accuracy value of results obtained in the first phase, and superior to the obtained results by Lasso method.

Another way to observe the effectiveness of the second phase is based on visualizing the obtained results. We first transform the ranking matrix $\mathbf{R}$ into position matrix $\mathbf{P}$. Element p_{ij} of $\mathbf{P}$ is the position index of the $j-th$ feature in the ranked list r_i, and it is equal to $p_{ij} = k$ where $r_{ik} = j$. By this way, the $j-th$ column of position matrix $\mathbf{P}$ is a list of position indices corresponding to the $j-th$ feature given by AE models. Then, we can visualize the position index of each ROI by using the boxplot [166]. Each boxplot will show the variation in the position index of a ROI, containing the values of minimum, maximum, mean, the first and the third quartiles, and outliers corresponding to the list of all possible position indices.

Fig. 3.11 depicts the position index of some ROIs before and after the second phase, respectively (see Appendix for full of ROIs). The blue dashed lines depict the mean

value of position index for every of 10 ROIs shown in the sub-figure. On the right column of the figure, we can observe that this line goes up from top to bottom which means that the average value of the position index is increasing. However, on the left column, this line seems to be unpredictable. Note that when comparing the contribution ranking between two ROIs, the one with the lower position index (or at the top of ranked list) will have a higher contribution ranking. It means that the right column appears to be sorted in ascending order of contribution which explains why the mean value of position index is increased. In the observation that each set of ranked features produced by a certain AE model in phase 1 gives a result unlike any other, as mentioned in Fig. 3.8 and Fig. 3.9. With the use of phase 2, we have picked out the good and stable ROIs that have been regularly voted by different AE models, as the first 10 ROIs in the right column of Fig. 3.11 has both a much smaller variance (shorter boxplot) and lower position index than the rest. It possibly explains why this phase can significantly improve the accuracy of classifiers when compared to average accuracy of the first phase and the results obtained by Lasso method.

From previous experiments, it shows that the proposed method is effective in providing a good criterion for ranking and selecting the most important ROIs. Its obtained results can help to significantly improve the accuracy of SVM classifiers. Furthermore, each phase of the proposed method also has a strong capability to select the most important ROIs which help effectively distinguish between AD and NC groups from PET images.

3.2.4 Conclusion

To address the issue related to randomly initialized weights for identifying the most important features to neural networks, in this study, we presented a simple supervised feature ranking method based on AutoEncoder and Connection Weights algorithm. Firstly, PET images were divided into regions of interest according to an anatomical atlas and mean intensity value of each region were computed. Next, we trained multi-AE models, the connection weights learned from these models were used to assess the contribution of each ROI to the network based on the Connection Weights algorithm. We then proposed a 2-phase feature ranking method for ranking and selecting the top k ROIs as candidate features for the SVM classifier. Experimental results on the ADNI dataset show that our method can significantly improve the accuracy of SVM classifier when compared to other methods such as: Fisher-Score, Lasso. Furthermore, each phase of the proposed method also has a capability to select the most relevant ROIs which help effectively distinguish between groups of subjects and better understanding about the diseases. While every single AE model in phase 1 is competitive to Lasso and can be used independently for feature ranking, phase 2 combines the best selection of ranked feature sets provided by phase 1 to obtain multi-view data analysis. It leads to a suggestion that instead of focusing on the complexity of an AE architecture, we can also take the advantage of multiple simple AE models to perform the same task. This method could be easily applied to any image dataset where a feature selection and feature ranking is needed. However, there

are two questions that need to be addressed when applying our method. The first one is how to choose the optimal value for Search Scope Value. The second is how to determine the minimum number of AE models that we need to train, or in the other terms can we reduce the number of models that we need to train. Therefore, for future work, it will be important to find the answers to these questions, as well as apply the proposed method to different datasets to better understand the limitations and strength of the proposed method.

3.3 Autoencoder-based feature ranking for MCI diagnosis

Recently, with developments of deep learning, there have been a lot of studies using neural networks for feature selections, particular the Autoencoder (AE). In [85], authors have incorporated autoencoder regression and $l_{2,1}$ regularization to select a subset of most useful features. In [58], authors proposed the Agnostic feature selection algorithm, combining an AE with structural regularizations to perform feature selection. In [1], authors have replaced the first hidden layer of AE with a "concrete selector" layer, which is the relaxation of a discrete distribution, and features are selected based on their probability of connecting to the nodes of the concrete layer. Instead of embedding feature selection in the learning process of AE, in a recent work [211], we proposed a 2-phase Autoencoder-based feature ranking (AEFR) method. In which, feature ranking is performed based on learning capabilities of different simple AE structures. While the obtained results were quite interesting, the proposed methods still need to be validated on other datasets. Moreover, it also has limitations in the second phase of the ranking algorithm, as we need to specify several parameters.

To overcome these limitations and verify the efficiency of the proposed method, in this work, we propose several improvements for our feature ranking method, and apply it to a more challenging task – MCI conversion diagnosis. More specifically, our contributions are:

- Improving the second phase of our ranking algorithm (parameter-free version).
- Improving the way for training AE, making the learning process more dependent on AE structures.
- Applying our method to MCI classification, which is a more challenging task.

The remainder of this paper is organized as follows. Section 3.3.1. describes the proposed method and the used dataset. Section 3.3.2. presents experimental results and comparisons. A conclusion and future work are then given in Section 7.6.

3.3.1 Proposed Method

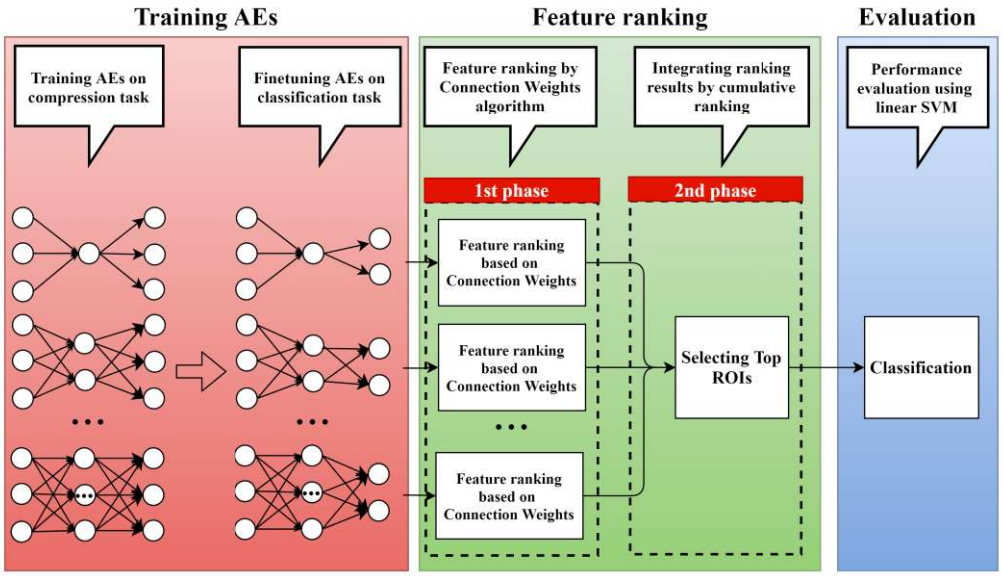

Figure 3.12: Flow chart of the proposed method

The flowchart of the proposed method is shown in Fig. 6.1. This flowchart can be divided into several blocks:

- Training Autoencoder models (in red).
- 2-phase feature ranking method (in green).
- Performance Evaluation (in blue).

In the following subsections, we introduce the proposed method in detail.

3.3.1.1 Dataset

Data used in the experiment is 18F-FDG PET images downloaded from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. In the ADNI dataset, participants can take several scans at different time points, the first time of scan refers to a baseline scan. Time points after baseline are considered as follow-up time. In this work, we only focus on predicting MCI conversion, data is selected based on following criteria:

- *sMCI - subjects diagnosed as MCI at the baseline and did not convert to AD in the follow-up period at least 24 months*;
- *pMCI - subjects diagnosed as MCI at the baseline and progressed to AD in the available scan time.*

After downloading the data, images are spatially normalized to MNI template using SPM12 [162] with $2 \times 2 \times 2 mm^3$ voxel size, $91 \times 109 \times 91$ tensor dimension. The intensity normalization is then performed through dividing each voxel intensity by the average value of the cerebellum region. After that, images are further smoothed by a Gaussian kernel with a full width at half maximum of 8 mm. All the procedures above are implemented with SPM12. After the preprocessing phase, we have a total 569 FDG-PET images for later experiments, including 209 pMCI subjects, and 360 sMCI subjects. The demographic and clinical information of subjects is provided in Table 6.1, in which MMSE stands for the Mini-Mental State Examination. After that, each subject image is segmented into 120 ROIs using AAL atlas [179]. Each region is then characterized by mean value. Since cerebellum is used for normalizing data, we will remove those regions from our data, after this step for each image, we obtain a vector with 94 features – corresponding to mean intensity of 94 ROIs.

Table 3.9: DEMOGRAPHIC AND CLINICAL INFORMATION OF SUBJECT

Characteristic	**sMCI**	**pMCI**
Number of subjects	360	209
Female/Male	153/207	87/122
Age (Mean ± std.)	71.73 ± 7.66	73.89 ± 6.88
MMSE (Mean ± std.)	28.20 ± 1.59	27.13 ± 1.71

3.3.1.2 2-phase feature ranking method

3.3.1.3 Auto-encoder training

An autoencoder [92] is a type of neural network used for learning efficient representation of unlabelled data, in compressed or sparse form. An AE usually consists of an input layer, an output layer and at least one hidden layer. The complexity of the AE network depends on the number of hidden layers and learning techniques. In our work, we only used simple AEs with one hidden layer. Instead of designing a single complex AE structure, we have focused on single-but-multiple AEs. This approach allows us to evaluate the stability of the features within minor changes in the network architecture during the feature ranking phase. Specifically, we trained 94 AE models with the number of hidden units varying from 1 to 94 on compression tasks. After that, we replaced the output layer of AEs by softmax layer and fine-tuned autoencoder models on the sMCI/pMCI classification task.

In comparison to our previous work [211], to ensure that the ranking results are more heavily dependent on the structure of AE models, we fixed the connection weights between the hidden-output layers during the AE training phase for the compression task. This constraint ensures that the learning process for compression tasks only happens

between input-hidden layers. Therefore, in the finetuning phase, we can replace the output layer of AEs with a softmax layer without losing any learned information from the compression phase. By this approach, we expect that the learned weights will be more dependent on AE structures, as well as more adapted for ranking problems.

3.3.1.4 Ranking methods

In order to rank all features, we use the same pipeline as proposed in [211]. After training AE models on compression tasks and finetuning them on MCI classification problems, we performed the first phase in our ranking algorithm. For each finetuned AE model, we used their learned connection weights and modified version of connection weight algorithm [211] to rank all features:

$$c_i = \sum_{j=1}^{p} |W_{ij}^{(1)}| \times \sum_{k=1}^{m} |W_{jk}^{(2)}| \tag{3.6}$$

where c_i is the contribution of the $i-th$ feature made to the output layer; n, p, m are the number of units in the input layer, hidden layer, and output layer, respectively. $W_{i,j}^{(1)}$ denotes the parameter of the connection between the $i-th$ neuron in the input layer and the $j-th$ neuron in the hidden layer, $W_{i,j}^{(2)}$ denotes the parameter of the connection between the $i-th$ neuron in the hidden layer and the $j-th$ neuron in the output layer. The higher value of c_i indicates the higher importance of the $i-th$ feature. After this step, for 94 AE models, we obtained a ranking matrix $\mathbf{R}$ in which $r_{i,j}$ corresponds to the $j-th$ most important feature ranked by finetuned model AE_i.

In order to combine results from multiple AEs into a single outcome, we continued with the second phase of our ranking algorithm. We have made several improvements to this phase compared to our previous work [211], resulting in a parameter-free version of our algorithm. Specifically, we have proposed a cumulative feature ranking algorithm, which is described in Algorithm 2. Given a ranking matrix $\mathbf{R}$ with the size of $n \times n$, in which n is the total number of features; sub matrices $\mathbf{R_j}$ correspond to the first $j-columns$ of $\mathbf{R}$. Firstly, we compute frequency matrix $\mathbf{F}$, where $f_{i,j}$ equals to frequency of $i-th$ feature in submatrix $\mathbf{R_j}$. After this step, we compute cumulative frequency for each feature, $C = \{c_i\}$, where $c_i = \sum_{j=0}^{n} f_{i,j}$. Then, we rank features in descending order according to their corresponding cumulative frequency. A higher value of cumulative frequency indicates the higher rank of the feature.

By using concepts of cumulative frequency, our algorithm has become a parameter-free algorithm. The complexity of the algorithm is less than $O(n^2)$, since we only need to compute the frequency of features in every submatrix R_j.

3.3.1.5 Experimental Design and Evaluation Metrics

To evaluate the effectiveness of our proposed method, we compared the performance between our method and other feature ranking methods on the ADNI dataset. We

Algorithme 2 : Parameter-free version of 2nd phase in ranking algorithm

Data : Matrix $R_{N\times N}$ corresponds to ranking order of N features given by N AutoEncoder models, sub matrix R_j contains first $j-columns$ of R

Result : Feature ranking list

begin

Compute frequency matrix $F_{N\times N} = \{(f_{i,j})\}$ where $f_{i,j}$ is frequency of $i^{th} - feature$ in sub matrix R_j

Compute cumulative frequency $C = \{c_i\}$ where $c_i = \sum_{j=1}^{N} f_{i,j}$

Ranking N features in descending order according to their cumulative frequency.

chose Fisher, T-Score, and Lasso [121] as comparative methods. Since, they are well-known and popular methods for feature ranking. After the ranking phase, for each ranking method, the top-k ROIs are selected as candidate features for the classification task. We choose linear Support Vector Machine (SVM) [48] as a classifier, since SVM is a well-known and widely used technique for AD diagnosing in previous studies. The performance of a the classifier is measured based on ACCuracy and Area Under Curve (AUC), since they are the most common and useful metrics to measure the overall performance of a binary classification model. The higher value of these metrics indicates better performance.

Concerning data allocation, we divide data randomly into two subsets corresponding to two phases: 50% for feature ranking phase, 50% for performance evaluation. Therefore, these two phases are independent learning processes. 5-folds cross validation is used for training and evaluating a classifier. To avoid bias, experiments are repeated for 50 times, and the reported results are averaged values of these different runtimes.

Finally, all experiments are implemented using scikit-learn [159] and Keras. For training Autoencoder we use the same settings as in our previous study [211]. For other experimental parameters, default values set by the libraries are used if not specified.

3.3.2 Results and Discussion

3.3.2.1 Results

Ranking methods significantly improve classifier performance (Table 3.10, Fig. 3.13). When comparing accuracy of linear SVM classifiers with and without using ranking methods, we observed that feature ranking methods give a better average accuracy. This is most noticeable in the range from 20 to 40 features, ranking methods improve ACC by about 8% on average compared to no ranking. Fisher and t-score gave very similar performance and they were the best methods in the range from 1 to 20 features, with an average ACC improvement of approximately 7% compared to no ranking. However, with more than 20 features, LASSO was the better one. Therefore, we chose

Table 3.10: Performance on sMCI/pMCI classification task

	Original		Fisher		t-score		LASSO		AEFR		Proposed method	
	ACC	AUC	ACC	AUC	ACC	AUC	ACC	AUC	ACC	AUC	ACC	AUC
k = 5	63.27	64.50	69.68	74.53	69.82	74.59	66.64	76.45	70.19	75.30	67.63	75.13
k = 10	63.27	65.00	70.93	74.76	70.95	74.80	68.60	77.12	70.92	76.33	69.79	75.35
k = 20	63.28	64.33	71.52	76.23	71.52	76.12	71.16	77.73	73.01	78.57	71.57	76.86
k = 30	63.31	62.38	71.77	76.64	71.80	76.58	71.93	77.79	73.93	79.21	73.02	78.12
k = 40	64.55	68.31	71.80	76.55	71.76	76.51	72.23	77.96	74.13	79.39	74.27	79.07
k = 50	65.73	71.57	71.76	76.56	71.92	76.60	72.56	77.90	74.11	79.32	74.29	79.24
k = 60	66.51	72.32	72.09	76.56	72.08	76.60	72.94	77.91	74.24	79.21	74.45	79.30
k = 70	73.12	78.28	72.33	76.64	72.38	76.68	73.21	78.08	74.32	79.23	74.37	79.18
k = 80	73.70	78.66	73.28	77.40	73.29	77.37	74.04	78.60	74.22	79.11	74.39	79.16
k = 90	73.83	78.60	74.19	78.70	74.16	78.70	74.26	78.83	74.20	78.98	74.36	79.00
k = 94	74.26	78.95	74.26	78.95	74.26	78.95	74.26	78.95	74.26	78.95	74.26	78.95

t-score and LASSO as the baseline when comparing cases with less than 20 features, and more than 20 features, respectively. Here, we expect that a good ranking method should give either similar or better performance to t-score, LASSO when using less than 20 features, and more than 20 features, respectively.

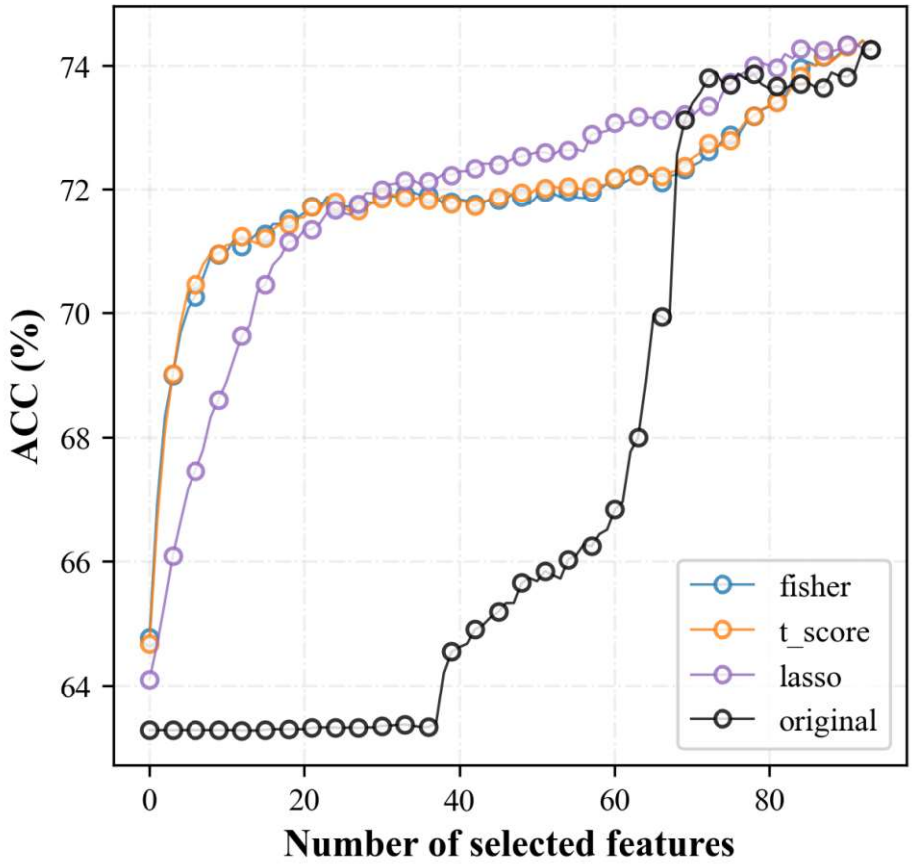

Figure 3.13: Comparing ACC of linear SVM classifier with and without ranking methods.

The AE-based ranking methods significantly improved classifier performance, as shown in Fig. 3.14 and Table 3.10. In many cases, ACC was improved by up to 10% compared to no ranking (e.g. the case with 30 features). AE-based ranking methods also gave better results when compared to the baseline methods, especially for AE models with fixed weights. Interestingly, after more than 30 features, the performance of classifiers seemed to reach a stable state with AE-based ranking methods, adding features almost did not change the performance much. It means that the AE-based ranking methods seem to have better convergence results when compared to baseline ones. Moreover, in the range from 40 to 60 features, there were many cases in which our proposed ranking methods gave better results than using either all 94 features or voxel-based approaches.

The ranking method with fixed AEs give a better performance compared to normal AEs, and outperform the baseline methods. More specifically, the top-1 feature selected by the fixed AE models gives a similar performance compared to t-score, and much better than lasso, normal AE method. The performance of fixed AE models is always approximately or better than the t-score in the range from 1 to 20 features. With more than 20 features the obtained results are much better when compared

to lasso. At the threshold of 30 features, the result seems to converge, faster than all the baseline methods. In other words, by using fixed AE models, we can achieve remarkable results with the number of features being only 1:3 of the total number of features, and approximately 1:1850 of features when compared to the voxel based approach.

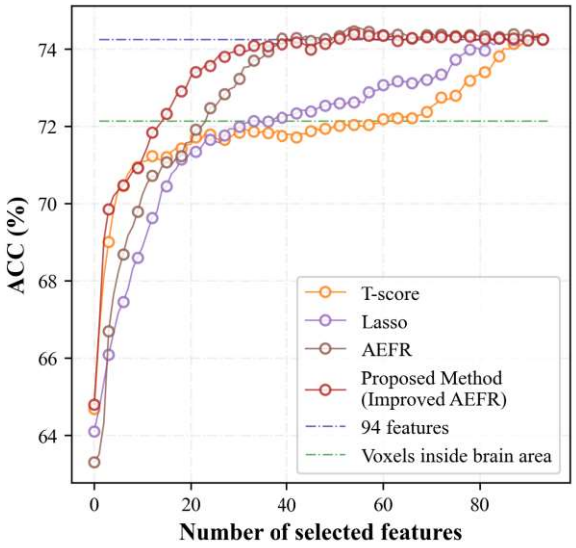

Figure 3.14: Comparing ACC of linear SVM classifier between our proposed methods and baselines.

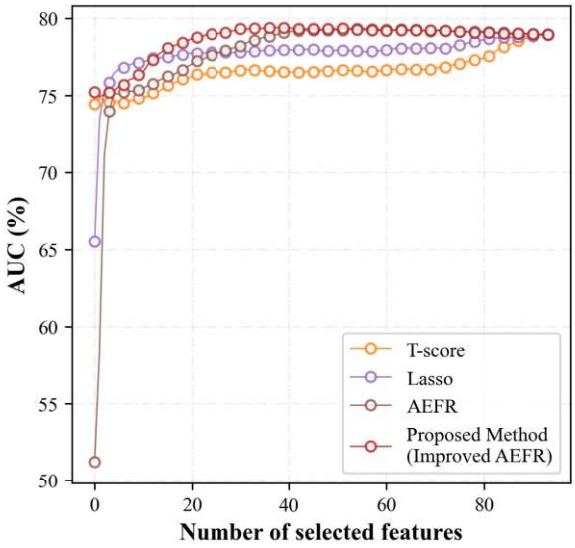

Figure 3.15: Comparing AUC of linear SVM classifier between our proposed methods and baselines.

Our proposed method not only significantly improves ACC but also improves AUC (Table 3.10 and Fig. 3.15). We can observe that ranking methods such as t-score and lasso improve AUC compared to original order by an average of 4.6%. The averaged differences in AUC between t-score and lasso did not differ too much. Meanwhile, the ranking methods using fixed AE and normal AE models, respectively, improved AUC approximately 1.5% and 1%, respectively, when compared to t-score. All these observations lead us to the conclusion that the AE-based ranking methods are suitable for feature selection on MCI classification tasks, and the ideas of fixed AE models can significantly improve the overall performance.

In order to explain the improvements between our proposed methods and the other, we considered the overlapped rate between obtained results by each ranking method. Fig. 3.16 compares the overlapped rate between different ranking methods with various numbers of selected features. We only focused on the range from 5 to 60 features. With the range from 5 to 40 features, the overlapped rate between lasso and t-score was around 35%, and these values increased to reach 60% within the range from 40 to 60 features. Obviously, these observations were corresponded to the differences in performances between two methods. With less than 15 features, the overlapped rates between fixed AE-based and baseline methods are above 60%

compared to t-score, and lower than 50% compared to lasso. As the number of features increased in the range from 1 to 15 features, we also observed that overlapped rates between fixed AE-based models and t-score decreased while the values between AE models and lasso increased. In other words, the ranking results obtained by fixed AE-based modes are quite similar to t-score with a small number of features, and become similar to lasso in the cases of much larger numbers of features. This observation confirmed our early expectation about a good ranking method, which should give either similar or better performance to t-score, lasso when using less than 20 features, and more than 20 features, respectively. This result also suggests that our fixed-AE method shows good qualities similar to those of both t-score and lasso.

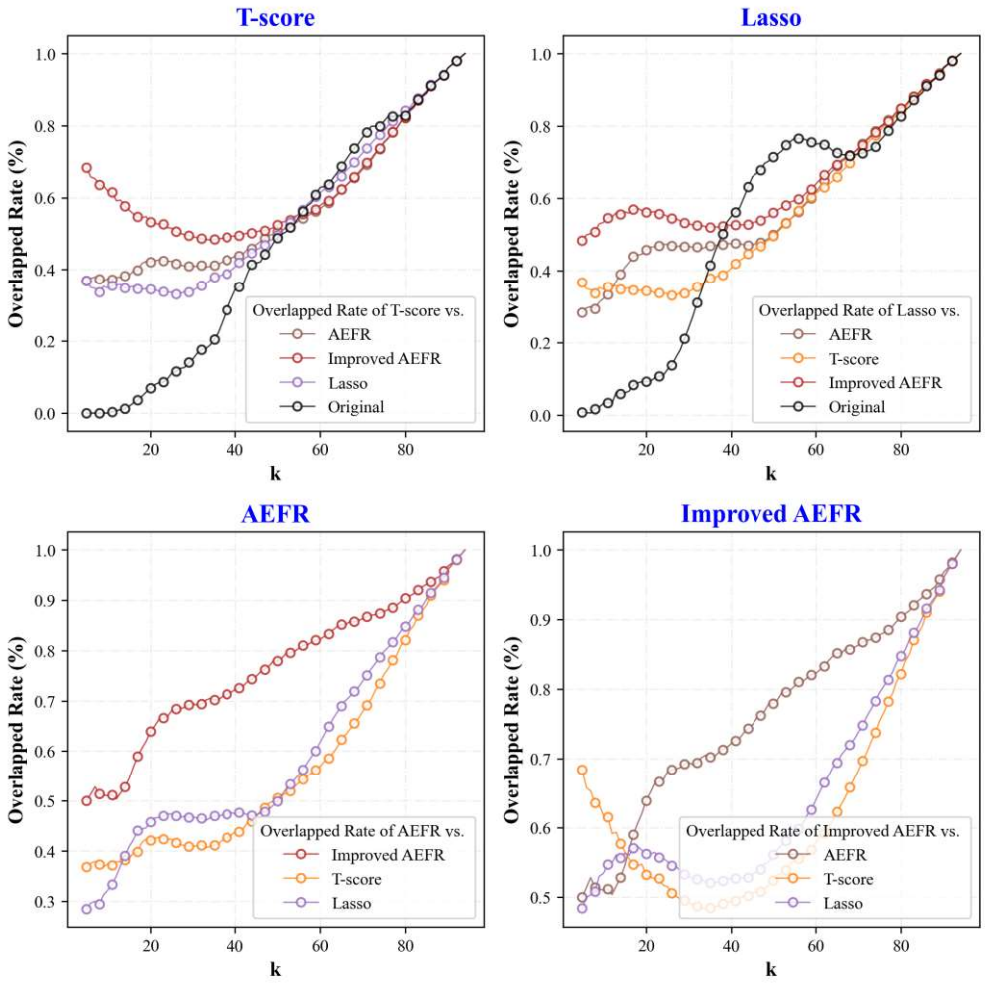

Figure 3.16: Overlapped rate between different ranking methods.

Moreover, the fixed AE method has its own capabilities which baseline methods do not have. Specifically, overlapped rates between the fixed AE method and the baselines are always less than 60% within the range from 20 to 60 features. Meanwhile, the fixed AE method gives a faster convergence results at the threshold of 30 features.

It suggests that the fixed-AE method could select some effective features that baselines are not able to do. These effective features not only improve the overall performance, but also contribute to stability and earlier convergence of results.

Fixing connection during training AE models significantly contributed to the improvement of obtained results. Comparing overlapped rates between fixed AE and normal AE models, we observed that the overlapped rate was around 50% in the range from 5 to 20 features. Meanwhile, with more than 20 features, the overlaps were always greater than 65% and continued to increase with the number of features. This observation not only explains the performance differences between two methods within the range from 1 to 20 features, but also explains why the performance between them is quite similar and converges together with more than 30 selected features. This result seems to support our early expectation of using fixed connection when training AEs for compression tasks. By using fixed connection weights, the learned weights will be more dependent on the AE structures and more adapted for ranking problems.

3.3.2.2 Discussion

In this study, we proposed two improvements to our previous AE-based feature ranking work: (i) fixing connection weights during AE training for compression tasks and (ii) enhancing the second phase of the ranking algorithm. Our proposed methods outperformed baseline methods such as t-score, fisher, and lasso. In particular, fixing connection weights approach significantly improves ACC and AUC, with early convergence and greater stability compared to both baseline methods and training AEs without fixing. This improvement can be explained as fixing connection approach restricts the learning process to happen only between input-hidden layers, enabling the connection weights between input-hidden layers to encode more information. As a result, more knowledge is learned which leading to improved performance on the MCI classification task and ranking results.

Fig. 3.17 and Table 3.11 shows the top-20 regions considered most effective in differentiating sMCI from pMCI, this obtained result is consistent with previous studies. More specifically, the top-20 regions mainly belong to the temporal and parietal lobes. These lobes are considered as the standard FDG-PET finding, which are highly correlated to AD pathology [24, 93]. Top-ROIs such as Temporal_Inf, Angular, Precuneus are also selected in previous studies of PAN on MCI conversion diagnosing [154], Tuan's study on AD [211]. Therefore, our results once again confirm the important role of these regions, and suggest that brain ROIs such as Angular, Angular, Precuneus can be considered as effective regions for tracking development and changes of AD in different stages.

In addition to similarity in the top-20 ROIs, there were also some significant differences between our obtained results with previous studies of PAN [154] and Tuan [211]. Despite sharing several common ROIs, our method gives a different ranking order for each ROI. Specifically, regions such as Angular_L, Angular_R, and Cingulate_Post_L are ranked in the top-3 by our proposed method. However, Angular_L and Angular_R are ranked quite low in PAN's study, while the Cingulate region does not appear in

Tuan's obtained results. It is noteworthy that many studies have reported a high association between these brain ROIs and the conversion of MCI to AD [152, 194, 21]. More specifically, with Angular_L and Angular_R belong to Parietal lobe, which are located in the posterior part of the inferior parietal lobule [187], in [93], authors have reported that hypometabolism in the inferior parietal lobules could be used to predict most reliably the progression from amnestic mild cognitive impairment to Alzheimer's disease. Regarding Cingulate, in a recent study, the authors pointed out a correlation between the Cingulate and apathy among MCI patients with and without apathy [56]. Meanwhile, apathy can be considered as a powerful risk factor for conversion to AD in MCI patients [178], especially when it occurs at the early stage of MCI [60]. Therefore, our obtained top-k regions seem consistent and coherent when compared to previous studies concerning MCI conversion.

In this study, we used multiple AEs for feature ranking. Experimental results suggest that we can use simple-but-multiple AE models to perform feature selection instead of focusing single-but-complex ones. By making several small changes in AE architecture (such as changing the number of neurons, fixing connection weights), and taking these changes into account for the feature ranking, the classifier's performance was significantly improved. The improvements can be explained by: (i) Training AE through two steps, the first one on compression task (unsupervised learning), the next one on classification task (supervised learning), which gives AEs the capability to filter out features that effectively represent both internal properties of the data and external properties related knowledge about classification problem. (ii) Subtle changes between AEs are enough to learn similar knowledge while creating variations for selecting important features with high stability. Together with previous results on ranking using AEs for AD [211], the obtained results suggest two things: (i) AEs could be used for selecting the most effective ROIs for distinguishing between HC, MCI and AD. (ii) instead of focusing on complex structures, we can also use simple-but-multiple AE models for feature selection with expectation that the obtained result is still very comparable to traditional feature ranking methods.

In comparison to our previous work, here, we proposed an enhancement to our ranking algorithm, resulting in a parameter-free version, making it easier to apply. While the obtained results are very promising, the approach has a limitation that may hinder its broad application - multiple AE models need to be trained, which raises concerns regarding computational resources. Nonetheless, due to the simple architecture of the AE models, consisting of only one hidden layer, the computational demand is low. Furthermore, fixing connection weights reduces the number of parameters required for training, resulting in a significant decrease in training costs. Moreover, AE models training can be performed in parallel, thereby reducing the training time. So, in general, our proposed method still has a broad applicability.

Table 3.11: Top 20 rois selected by proposed method

	Top ROIs
1	Angular_L
2	Angular_R
3	Cingulate_Post_L
4	Amygdala_L
5	Temporal_Inf_L
6	Precuneus_R
7	Parietal_Inf_R
8	Parietal_Inf_L
9	Putamen_R
10	ParaHippocampal_L
11	Precuneus_L
12	Pallidum_L
13	Cingulate_Post_R
14	Putamen_L
15	Postcentral_R
16	Paracentral_Lobule_L
17	ParaHippocampal_R
18	Temporal_Inf_R
19	Paracentral_Lobule_R
20	Postcentral_L

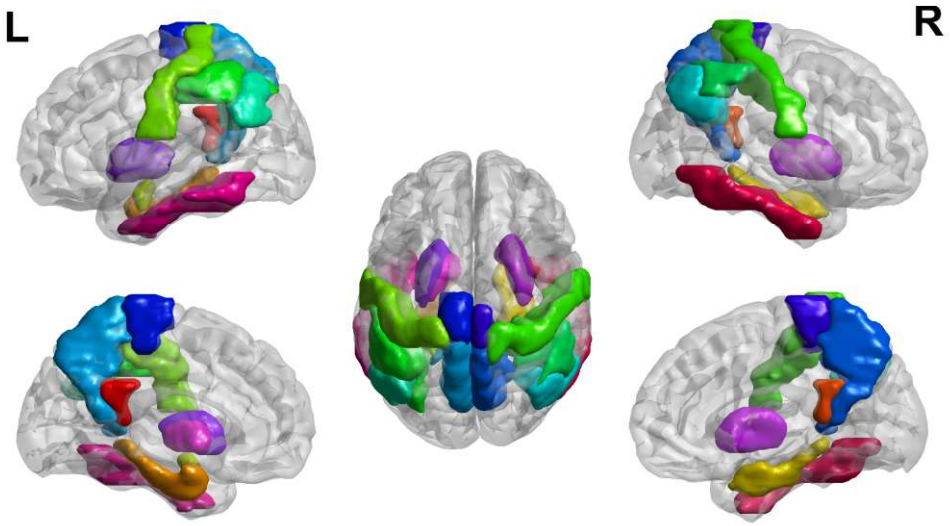

Figure 3.17: An illustration for top-20 ROIs selected by our proposed method (with fixing connections)

3.3.3 Conclusions and Future work

In this paper, we have made some improvements to our previous work on AE-based feature ranking. Specifically, we fixed the connection weights between hidden-output layers when training AEs on compression tasks; and improved the second phase of the ranking algorithm. We applied our method to a very challenging task – MCI conversion prediction. Experimental results on the ADNI dataset showed that our method outperforms baseline methods such as t-score and Lasso. For feature selection, these results suggest that instead of focusing on designing a single complex AE architecture, we can also take advantage of simple-but-multiple AE models to perform the same task. Our proposed method could be easily applied to a dataset where a feature selection is needed. For future work, it is important to apply our proposed method to different datasets to better understand the limitations and strengths of the method. Additionally, considering the influence of other factors such as dropout and activation functions on the obtained results would be an interesting research direction.

3.4 Chapter Conclusion

This chapter presented Autoencoder-based frameworks for ranking and selecting the most crucial features highly affected by AD. The experiments have demonstrated the effectiveness of our proposed method. Several conclusions can be drawn from the observations made in these experiments:

- AE can effectively select vital features for diagnosing AD. The selected ROIs can serve as indicative regions for diagnosing AD.

- Instead of designing complex AE architectures, utilizing simple yet multiple AE models can potentially achieve performance equal to or better than traditional feature ranking methods.
- Feature ranking appears to enhance classifier performance. Some regions exhibit continuous changes throughout the course of AD, while others change at specific intervals, indicating the presence of stable regions.

In this chapter, we used AAL with a fixed number of regions for analysis, yielding several interesting results. However, it remains crucial to determine whether the choice of atlases affects the final outcomes. Should we utilize atlases at multiple scales? Can we develop adapted atlases more suitable for brain analysis? The subsequent chapter will delve into finding answers to these questions.

Chapter 4

Adapted Brain Atlas

Summary

4.1 Introduction . . . 89
4.2 C-Atlas . . . 90
4.2.1 Introduction . . . 90
4.2.2 Proposed method . . . 91
4.2.2.1 Dataset . . . 91
4.2.2.2 Brain Mapping Construction . . . 93
4.2.3 Proposed method . . . 93
4.2.3.1 SVM-based phase . . . 93
4.2.3.2 SLIC-based phase . . . 94
4.2.3.3 Experimental Design and Evaluation Metrics . . . 95
4.2.4 RESULTS AND DISCUSSION . . . 95
4.2.4.1 Feature Extraction . . . 95
4.2.4.2 Feature ranking . . . 97
4.2.4.3 Discussion . . . 98
4.2.5 CONCLUSIONS AND FUTURE WORK . . . 100
4.3 Chapter Conclusion . . . 101

4.1 Introduction

The previous chapter demonstrated that selecting the most crucial ROIs significantly enhances classifier performance. In this context, each ROI is determined by its location and the voxels associated with it. Typically, predefined atlases like AAL, Schaefer, and Hammer atlases are commonly used to define ROIs. Clearly, the choice of atlases greatly impacts the final analysis results. These atlases vary in the number and location of regions, leading to distinct effects on the analysis outcomes. While many atlases operate at a single scale, several are designed at multiple scales (Schaefer atlas). In the previous chapter, we solely employed the AAL atlas at a single scale. However, our brain is a complex system wherein neural connections form brain regions, and these regions collectively shape human behavior. Therefore, understanding our brain as a

complex, multiscale system is crucial. Studying the brain at multiple scales becomes more significant due to this inherent complexity.

This chapter addresses the significance of studying the brain at multiple scales using brain mapping. We focus on two key questions:

- Does scale matter? Can increasing scales enhance diagnostic performance?
- Can we develop adaptive brain mapping tailored to specific brain disorders?

The first question directly relates to studying the brain at multiple scales. The second question arises from the observation that most predefined atlases are generalized for brain studies. However, different brain disorders affect the brain in distinct ways. Consequently, predefined atlases might not be optimally effective, potentially dividing the brain into excessively small or large areas. Therefore, exploring the development of brain mapping tailored to specific brain disorders is intriguing. Such brain mapping would not only delineate distinct brain regions but also provide information emphasizing each area's role in the course of the disease.

4.2 C-Atlas

4.2.1 Introduction

Alzheimer Disease (AD) is a major type of neurodegenerative brain disease in elderly people. In an estimation of [157], fifty million people worldwide are living with dementia in 2018, and this number will triple in 2050. Recent years, there has been a variety of studies concerning how to distinguish AD from non-AD group (or Normal Control – NC), using machine learning and FDG-PET, a nuclear medicine functional imaging that uses 18F-FDG as a radiotracer [5]. With the ability to search for specific characteristics of the disease in vivo, PET is preferably used, more than Magnetic Resonance Imaging (MRI) or other similar scanning technology, for early AD detection.

With FDG-PET images, previous studies used either voxel-based or region-based approach. The first approach uses all voxels of the whole brain as learning features, as studied in [37, 90], while the second one segments the brain into different region of interest (ROI) and then features are extracted from these regions, as studied in [149, 78]. The region-based approach is preferable because it can help reduce the dimension of data and the effect of noise as well as facilitate image analysis. One of the ways to segment brain image into regions is to use predefined brain atlases such as Anatomical Automatic Labeling (AAL) [215], LONI Probabilistic Brain Atlas (LPBA40) [190], Talairach atlas [117]. These atlases are built strongly on anatomical landmarks, cytoarchitectonic information. Despite having a lot of neuro-biological meanings, these atlases may not accurately reflect information about the structure and function of the brain when specifically applied to a certain type of brain disease such as AD [12]. Meanwhile, building a new brain atlas adapted to specific subjects or certain diseases is an approach of interest, where the most relevant information

associated with the disease, together with functional changes from subject to subject within and between stages, is captured.

Regarding the study of AD, it is essential to develop a brain mapping suitable for the disease. In our literature review, there are a lot of studies related to building a new brain mapping. Some whose build is based on random criteria [84], on structural criteria [246] or on information about connectivity between regions [51] with a lot of different clustering techniques, such as: k-means algorithms, Gaussian mixture models, spectral clustering and more as extensively reviewed in [12]. Despite interesting results of these studies, we will encounter the following compatibility problems in our scope of PET images and AD. Firstly, these atlases are mainly built from MRI images of healthy control subjects [64], many of which specifically focused on functional MRI (fMRI) images in the resting state of the patient (rs-fMRI) [51, 209, 23]. This means that using these atlases might be inappropriate to extract the information from damaged brain regions. Moreover, there is a significant difference between PET and MRI images in terms of how data linked to medical information is acquired. As a result, using brain mappings which are learned from MRI images to extract features from PET images may not be suitable. Secondly, with the condition that data comes from healthy control subjects, previous studies are not possible to take full advantage of the information regarding the disease progression (e.g., patients with and without diseases). We hypothesize that incorporating the information from different patient groups is extremely important for learning adaptive brain mapping for specific kind of disease. Therefore, we propose a method to build a functional brain atlas from PET images, using Simple Linear Iterative Clustering (SLIC) [2] algorithm as a segmentation method and Support Vector Machine (SVM) [49] classifier as a guiding criterion. By doing this way, we expect that the learned atlas will be more suitable for extracting features from PET images when compared to manual and predefined atlases. The effectiveness evaluation is done on Alzheimer's Disease Neuroimaging Initiative dataset for both feature extraction and feature selection (ranking) aspects. The performance of C-Atlas is compared with AAL, LPBA40 (single scale) and of Schaefer [185] atlases (multiple scales).

The remaining of this paper is organized as follows. Section II. describes the proposed method and the used dataset. Section III. presents experimental results and comparisons. A conclusion and future work are then given in Section IV.

4.2.2 Proposed method

4.2.2.1 Dataset

Data used in the experiment is 18F-FDG PET images downloaded from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). After the downloading process, images are then spatially normalized to Montreal Neurological Institute (MNI) template with the shape of $91 \times 109 \times 91$. Next, the intensity normalization is then performed by dividing each voxel intensity by the average value of the cerebellum region. Finally, images are smoothed by a Gaussian kernel with a full width at half

maximum of 8 mm. More detail about data selection and data preprocessing can be found in our previous published work [211]. After preprocessing phase, we have a total 479 FDG-PET images for later experiments, including 237 AD subjects, and 242 NC subjects. The demographic and clinical information of subjects is provided in Table 4.1, in which MMSE stands for the Mini-Mental State Examination.

Table 4.1: Demographic and clinical information of subject

Characteristic	**CN**	**AD**
Number of subjects	242	237
Female/Male	122/120	97/140
Age (Mean ± std.)	73.66 ± 5.66	75.00 ± 7.91
MMSE (Mean ± std.)	29.03 ± 1.20	23.19 ± 2.12

4.2.2.2 Brain Mapping Construction

4.2.3 Proposed method

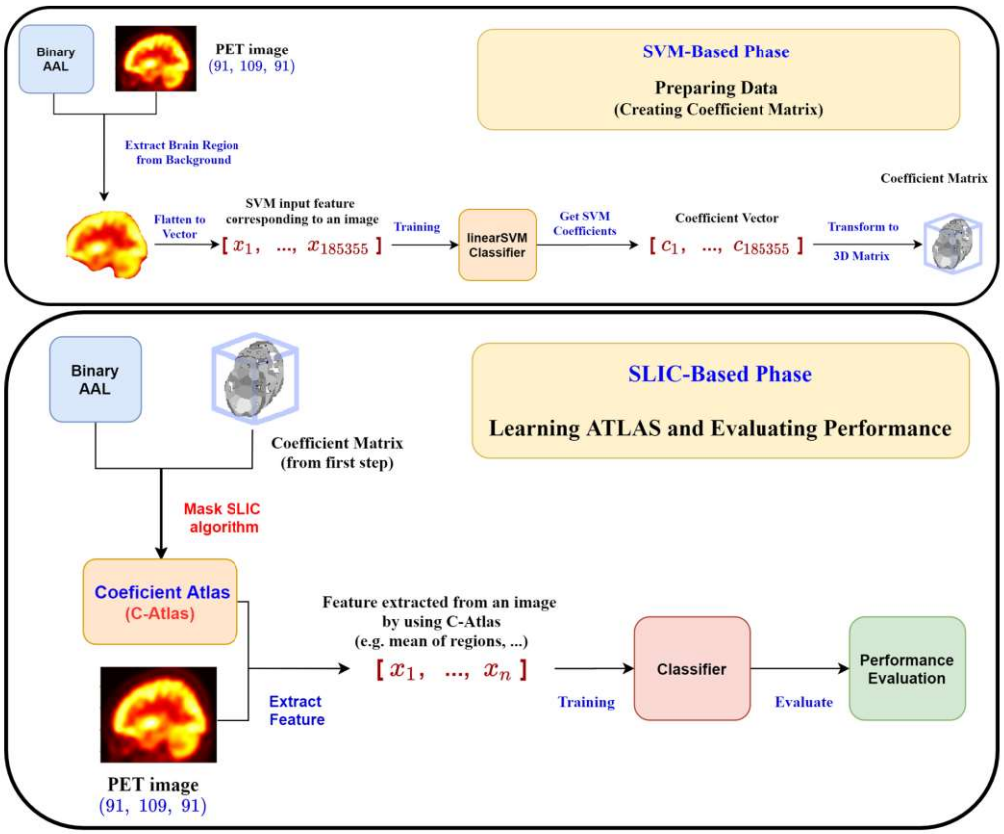

Figure 4.1: The flowchart of the proposed method, including 2 phases: SVM-based (top) and SLIC-based (bottom).

The process of learning new brain mapping atlas can be divided into two phases, corresponding to two blocks in Fig. 4.1. The first block denoted as SVM-based phase corresponds to pre-processing steps (data preparation for atlas learning). The second block denoted as SLIC-based phase corresponds to the learning process of C-Atlas and evaluating its performance. In the following subsections, we introduce the proposed method in detail.

4.2.3.1 SVM-based phase

In the first phase, we prepare input for learning atlas. In order to do this, we use PET images and their corresponding labels (AD/NC) for training a linear SVM classifier. Here, we choose the linear SVM classifier for two reasons: Firstly, it is simple enough and can be considered as good choice when the data is of high dimensionality (approximately 10^6, in our case). Secondly, the weights learned by linear SVM can be

used to decide the relevance of each feature [83]. This property is critically important for the next phase of the proposed method. When training a linear SVM classifier, we only focus on voxels corresponding to brain area. Obviously, training SVM with the background (area outside the brain) is ineffective and computationally expensive. To select only voxels inside brain area, we use a binary version of AAL (bAAL). Then, we flatten the 3D brain region into a vector, since the linear SVM classifier only takes input in the form of a vector. From each PET image in 3D, after these steps, we obtain a vector with approximately 2×10^5 (more specifically, 185355) elements, whose size is matched with the size of brain region in bAAL.

After SVM training process, we obtain a coefficient vector c with weight c_i corresponding to the $i-th$ feature of the SVM classifier, indicating the relevance of this feature to the classifier. Since, the value of the coefficient can be negative or positive, in order to facilitate later experiments, we normalize coefficient vector into the range from 0 and 1 using min-max normalization. After that coefficient vector is reshape into original size of input image (voxels outside brain area set to 0).

4.2.3.2 SLIC-based phase

Coefficient matrix obtained from the earlier phase is then used as the input to construct a new brain mapping. While most of the previous studies have directly input image dataset into segmentation algorithms, in our work, the coefficients learned by SVM are fed into them. We believe that using a classifier as a guiding criterion can help us take advantage of data labels, especially in case where data can be divided into multiple groups, each of which has its own label and significant structural differences. In addition, by using classifier's weights as a criterion for building atlas, we also believe that the learned atlas may be more suitable for extracting features fed into the final disease classifier. To the best our our knowledge, there has not yet been studies which used classifier weights as input for learning a new brain mapping.

For segmentation algorithm, we use SLIC algorithm [2] due to several reasons. Firstly, SLIC – a variation of k-means algorithm, is one of the most simple and efficient segmentation algorithms. Secondly, in the scope of our study, we only focus on voxels in brain area. All voxels in background should be excluded from the calculation of segmentation algorithm. Meanwhile, there is an extension of SLIC that allows us to perform the segmentation task only on certain regions of interest (e.g., brain area in our case) instead of the whole image, called maskSLIC [104]. In addition, comparing to other segmentation algorithms, SLIC has very few parameters that need to be optimized. More specifically, there are three important parameters that should be concerned: number of expected regions, shape compactness, number of iterations. Region number is an open parameter that will vary at observation desire. Shape compactness can be tuned up automatically with the mode slic0. Iteration number, despite being set as 10 by default for most problems, is set as 200 as the optimized value in our experiment tests.

4.2.3.3 Experimental Design and Evaluation Metrics

After learning the new brain mapping, we need to evaluate its effectiveness in the problem of AD diagnosis. In order to do so, we design experiments on ADNI dataset for AD/NC classification task. We use the learned atlas to extract informative features from PET images and use these values as input for linear SVM classifier. Given by the classifier, TP, TN, FP and FN are true positive, true negative, false positive and true negative, respectively. The performance is then measured based on ACCuracy, SPEcificity, SENsitivity and Area Under Curve (AUC) metrics where:

$$\begin{aligned} ACC &= \frac{TP+TN}{TP+TN+FP+FN}, \\ SEN = \frac{TP}{TP+FN}, SPE &= \frac{TN}{TN+FP} \end{aligned} \tag{4.1}$$

AUC is the Area Under the Curve ROC (Receiver Operating Characteristics) determined by SEN and $1-SPE$. The higher value of these metrics indicates the better performance.

Concerning data allocation, we use different dataset for each phase of constructing and evaluating the C-Atlas. Data is randomly divided into two subsets corresponding to two phases: 50% for learning atlases, 50% for evaluating these atlases. Therefore, these two phases are independent learning processes. Moreover, the nature of data learning in the first phase is different from the second. That is, while the construction process relies on the intensity of brain voxels to learn coefficient weights, and uses these weight values to define brain regions; the latter considers the mean intensity of previously defined regions applied to actual brain image. By doing this way, the two phases of proposed methods are confirmed as independent; thus, avoid the double dipping. Finally, all experiments are implemented using scikit-learn [159] and scikit-image [218] libraries. For experimental parameters, default values set by the libraries are used if not specified. To avoid bias, experiments are repeated for 100 times, and the reported results are averaged values of these different runtimes.

4.2.4 RESULTS AND DISCUSSION

4.2.4.1 Feature Extraction

In order to evaluate the effectiveness of C-Atlas on AD classification task, we use C-Atlas to extract informative features from PET images and use these values as candidate features for training AD classifiers. More specifically, PET image is segmented into regions using the C-atlas. We then compute the mean value of each brain region and use these values as input features for training linear SVM classifier. 10-fold cross-validation is used for each experiment. Here, we mainly focus on region-mean values since they can be considered as a good indicator to observe the difference between AD and NC subject in PET images. Normally, mean values of abnormal regions will be lower for AD subjects. Furthermore, region mean intensity is also a popular feature used in previous studies [149, 78].

Table 4.2: Performance on AD/NC classification task

	# Features	**ACC**	**SEN**	**SPE**	**AUC**
Full Voxel	902629	89.81 ± 2.09	87.67 ± 2.69	91.84 ± 2.27	95.91 ± 1.26
Voxel in Brain Area	185355	91.51 ± 1.83	89.96 ± 2.45	92.95 ± 1.90	96.65 ± 1.08
LPBA40	56	87.55 ± 1.55	84.50 ± 2.63	90.50 ± 2.05	94.22 ± 1.19
AAL	120	88.13 ± 1.70	85.42 ± 2.67	90.78 ± 1.98	95.04 ± 1.24
Schaefer (100)	100	90.22 ± 1.46	87.27 ± 2.58	93.06 ± 1.72	95.80 ± 1.12
Schaefer (200)	200	90.99 ± 1.50	88.40 ± 2.60	93.47 ± 1.69	96.34 ± 1.06
Schaefer (500)	500	91.12 ± 1.88	89.49 ± 2.62	92.66 ± 1.90	96.40 ± 1.20
Schaefer (1000)	1000	90.49 ± 2.13	89.11 ± 2.57	91.77 ± 2.27	95.94 ± 1.50
C-Atlas (50)	50	86.82 ± 1.66	82.66 ± 2.79	90.88 ± 1.91	93.88 ± 1.20
C-Atlas (100)	100	89.09 ± 1.50	85.34 ± 2.57	92.74 ± 1.74	95.61 ± 1.09
C-Atlas (120)	120	89.55 ± 1.50	85.91 ± 2.72	93.08 ± 1.77	95.90 ± 1.05
C-Atlas (200)	200	90.87 ± 1.44	88.12 ± 2.45	93.53 ± 1.68	96.50 ± 1.00
C-Atlas (500)	**500**	**91.83 ± 1.57**	**90.04 ± 2.33**	**93.54 ± 1.60**	**96.84 ± 1.01**
C-Atlas (1000)	1000	91.35 ± 1.75	89.79 ± 2.30	92.82 ± 1.91	96.66 ± 1.12

We then compare the performance of C-Alas at different scales with the results of voxel-based approach, fixed-scale predefined atlases (AAL, LPBA40); and Schaefer – a multi-scale atlas. From Table 4.2, we can observe that C-Atlases with more than 100 regions always give a better performance than predefined atlases, the most obvious improvements are at ACC (approximately 1% or more) and SPE (approximately 1.5% or more). C-Atlases with more than 200 regions, give a better performance than full voxel case, and, it is quite similar or even better when compared to all-voxel while the number of features is just approximately $500/902629 \approx 1/1800$ and $500/185355 \approx 1/370$ less, respectively. When compared to Schaefer atlas with matched number of regions (Fig. 4.2), most of the cases, C-Atlas gives a better performance for all evaluation metrics, especially for ACC, SPE and AUC. It means that the new learned C-atlas seems to be more effective when compared to predefined atlases. At least, in the case of using them to extract features from PET, the results are significantly improved.

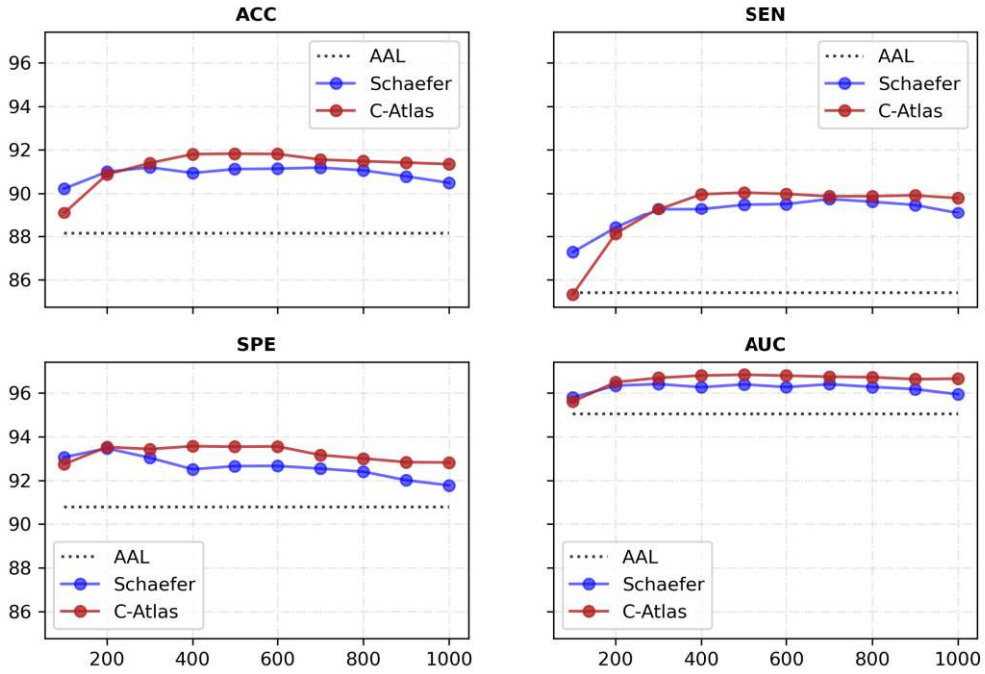

Figure 4.2: Performance comparison between C-Atlas and Schaefer.

4.2.4.2 Feature ranking

We have used C-Atlas for feature extraction. In comparison to predefined atlases, previous results suggested that C-Alases are more suitable for AD diagnosis on PET images. However, the ability of C-Atlas is more than that. We can also use C-Atlas for feature ranking. We know that the coefficient vector from linear SVM can be used for deciding relevance of each feature to the classifier. Based on this concept, we can rank each region R_i of C-Atlas by computing its coefficients values c_i which equals to sum of SVM coefficients of all voxels in its corresponding region Eq. 4.2. After computing total coefficient values for each region, we can rank them together. The higher value of a coefficient, the higher relevance of a region to a classifier. Fig. 4.3 is a boxplot comparing region values between NC and AD groups for top 9-selected regions ranked by the coefficient of C-Atlas. We can observe that, these regions show a clear difference in region values between the two groups. This result suggests that C-Atlas can be also used for feature ranking.

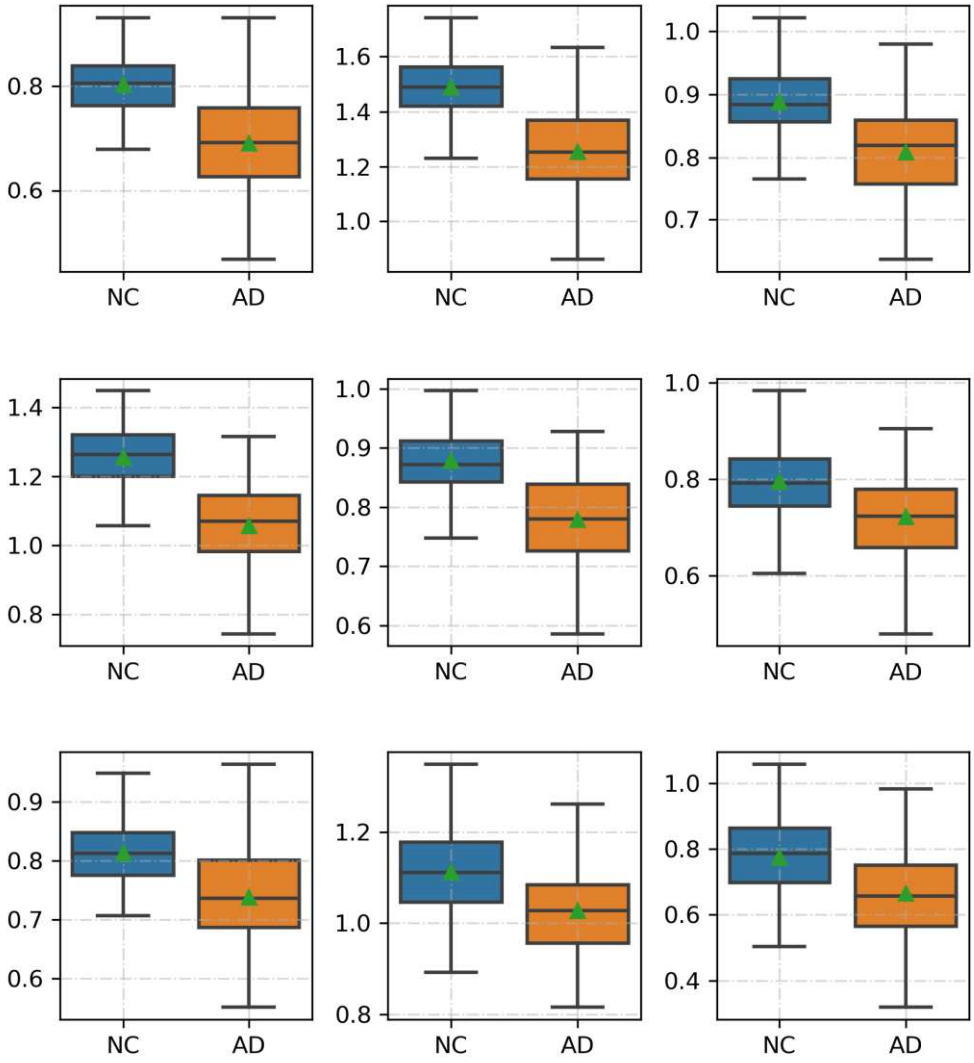

Figure 4.3: Boxplot for top-9 ROIs of C-Atlas (200), comparing distribution of data between NC and AD group.

$$c_i = \sum_{voxel_j \in R_i} w_j \tag{4.2}$$

4.2.4.3 Discussion

When using C-Atlas for AD classification problem, C-Atlas gives better performance than predefined atlases. In several cases, the improvement is not significant. However, due to characteristic of a data-driven atlas, C-Atlas can be enhanced. Since we used just a small set of data in this study, when more data is available, we can expect that C-Atlas can be improved and therefore maybe outperform predefined atlases. By inheriting the characteristics of linear SVM coefficient, in addition to the ability of

dividing the brain into separated regions, C-Atlas also has the ability to rank these regions according to their relevance for the AD classifier. The observations from boxplot (Fig. 4.3), has shown that regions ranked by C-Atlas can be considered as effective regions for diagnosis and tracking changes of the disease. Thus, compared with predefined atlases, C-Atlas has an added advantage in ranking the relevance of each region to the disease. Fig. 4.4 is an illustration of C-Atlas with 200 regions. When compared with AAL (120 regions), we can find some brain regions that seem not to differ too much between two atlases (red circle). However, we can observe that there are regions of the AAL that are subdivided in the C-Atlas (black circle), meanwhile there is a region in C-Atlas which is a combination of several regions in AAL (blue circle). Comparing with the differences in performance for AD classification task, those structural difference suggest that regions in the C-Atlas seem to be "adapted" to be more suitable for AD diagnosis when compared to AAL.

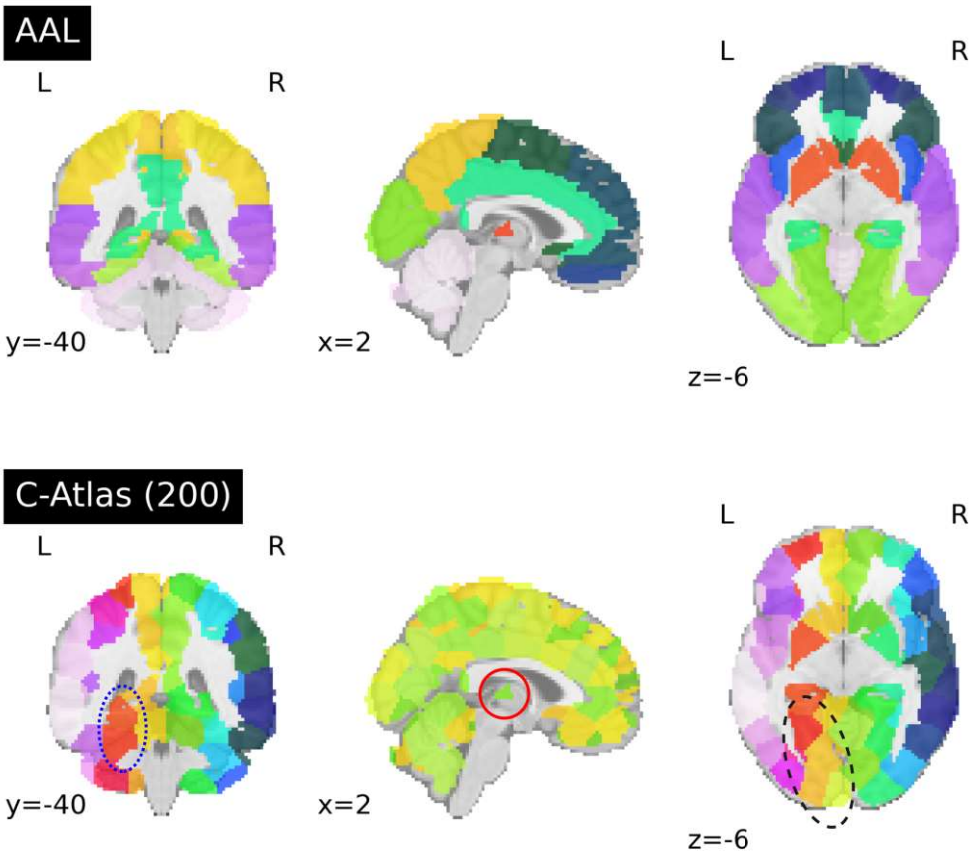

Figure 4.4: Comparison visualization between AAL (120 regions) and C-Atlas (200 regions).

When comparing the location of each region in C-Atlas with AAL, we can observe that these top ranked regions by C-Atlas (which can be considered as important indicators for AD), respectively belong to brain ROIs such as: Parietal, Frontal ROIs of

Table 4.3: Location of C-ATLAS (200) Top-Ranked Regions comparing to AAL

top-10 ROIs	**Automated Anatomical Labeling**
1	Parietal_Inf_R, Parietal_Sup_R, Angular_R
2	Parietal_Sup_R, Precuneus_R
3	Frontal_Sup_Medial_R, Frontal_Sup_2_R
4	Cingulate_Ant_R, Frontal_Sup_Medial_R, Frontal_-Med_Orb_R, Frontal_Sup_2_R
5	Cuneus_L, Occipital_Sup_L
6	Occipital_Sup_L, Occipital_Mid_L
7	Hippocampus_L, ParaHippocampal_L, Thalamus_L, Fusiform_L
8	Precentral_L, Postcentral_L
9	Parietal_Inf_L, Postcentral_L, Parietal_Sup_L
10	Temporal_Inf_L, Fusiform_L

AAL. These ROIs are considered as one of ROIs most affected by AD [25]. It seems that the top-selected ROIs by C-Atlas are consistent with the results of previous studies. Moreover, from Table 4.3, we can also observe that the top-ranked ROIs of C-Atlas are the combinations of sub-ROIs of AAL. It seems that C-Atlas selects the most important part of AAL and combining them into a larger ROI, which can explain why C-Atlas is more effective when comparing to AAL for AD diagnosis. For example, the top-1 ROI of C-Atlas is a combination of smaller parts from Parietal_Inf_R, Parietal_Sup_R, Angular_R region of AAL. These regions are, according to recent studies, considered to be highly affected by AD [211]. Comparing to AAL in terms of structure, above observations suggests that the structure of C-Atlas is adapted in size, and its top-selected-regions is a combination of several important regions which are heavily influenced by AD.

4.2.5 CONCLUSIONS AND FUTURE WORK

In this paper, we have proposed a 2-phase method for building a new brain mapping (C-Atlas) from PET images. We evaluated the effectiveness of C-Atlas for AD classification task. From experiments, it shows that our C-atlases are more effective than predefined atlases in terms of both feature extraction as well as feature selection. This can be explained by the design of C-Atlas more adapted to AD in order to extract and select the most informative features from PET images. The obtained results suggest that it is possible to learn brain atlases which are adapted to a specific kind of disease, and these atlases are promising to help us better understand the disease. Our method could be applied to learn brain mapping with varying number of ROIs, so it could be a useful aid for analyzing brain network at multi-scales. Moreover, not limited in healthy

control subjects, our method also considers data from different group of subjects; therefore, it provides a framework for learning and comparing brain mapping between different type of diseases.

In this work, we only focus on PET images, and we use SVM classifier as a criterion and SLIC as segmentation algorithm for constructing the new atlas. In the future, considering working with other modalities, such as MRI, fMRI; conducting some experiments with other types of learning methods such as deep learning-based or graph-based can be seen as interesting approaches. Moreover, currently, we are also investigating our proposed method for other types of brain disorders for evaluating the effectiveness of our method.

4.3 Chapter Conclusion

This chapter introduces a method for constructing adaptive brain mapping from FDG-PET images. The experiments conducted revealed the following key points:

- Studying the brain at multiple scales is crucial; scaling the mapping demonstrates a corresponding change in performance. Considering multiple scales is expected to notably enhance classifier performance.
- Developing an adaptive brain mapping tailored to brain disorders, specifically for FDG-PET and AD, is achievable.
- This work follows a data-driven approach to constructing a brain mapping, introducing the concept of "coefficient brain mapping." This concept not only delineates brain regions but also provides coefficients measuring the importance of these regions and their roles in brain disorders.

Together with the previous chapter, this section underscores the importance of carefully selecting methods to segment the brain into regions, investigating the contribution of each region to brain disorders, and emphasizing the necessity for multi-scale studies. By incorporating various information about atlases and ranking the importance of features, significant improvements can be made in classifier performance. Simultaneously, this approach identifies the most crucial regions, serving as indicators for diseases. This aids in tracking changes during the course of the disease and targeting treatment interventions.

Chapter 5

Metabolic Network: A Short Review

Summary

5.1 Introduction . . . 104
5.2 Methodology . . . 105
5.2.1 Research Questions . . . 105
5.2.2 Search Strategy and Eligibility Criteria . . . 106
5.3 Basic Concepts . . . 106
5.3.1 Graph Theory . . . 106
5.3.2 Construction of Metabolic Connectivity network using FDG-PET 107
5.3.3 Computing Graph Measures . . . 108
5.3.4 Most common graph metrics . . . 109
5.3.4.1 Network Degree and Network Density . . . 109
5.3.4.2 Shortest path length and Characteristic Path Length . . 110
5.3.4.3 Network efficiency . . . 110
5.3.4.4 Number of Triangles and Clustering coefficient . . . 110
5.3.4.5 Nodal centrality . . . 111
5.3.4.6 Small-world . . . 111
5.4 Results . . . 112
5.4.1 Literature Search . . . 112
5.4.2 Study Characteristics . . . 113
5.4.3 General Overview . . . 113
5.5 Discussion . . . 117
5.5.1 Healthy Control-based Metabolic Network Analysis . . . 117
5.5.2 Metabolic Network Patterns in Brain Disorders . . . 118
5.6 Challenges and Future Directions . . . 123
5.6.1 From Data Perspective . . . 123
5.6.2 Building and Analyzing brain metabolic networks . . . 123
5.6.3 Neural Network as a tool for exploring brain connectivity . . . 125
5.6.4 Clinical Intervention . . . 125
5.7 Conclusion . . . 125

5.8 Chapter Conclusion 126

Up to this point, we've discussed the significance of key brain regions throughout the course of AD and their impact on classifier performance. However, it's crucial to note that everything discussed so far hasn't taken into consideration the interactions among multiple regions. In reality, during the progression of the disease, several regions can be affected simultaneously. The magnitude of their influence varies depending on their locations, disease stages, and individual patient conditions (such as gender, education, clinical history, etc.). Considering interactions between different brain regions and variations among subjects is pivotal for more accurate diagnosis, marking a crucial step towards personalized treatments. In the upcoming chapters, we'll explore graphs as powerful tools for studying interaction patterns among various brain regions.

This chapter is dedicated to graphs and their applications in analyzing brain disorders. Initially, we'll provide a brief explanation of what a graph is, highlighting its importance in studying brain disorders. Additionally, we'll underscore several applications of graph analysis in studying metabolic brain networks. The chapter also serves as a foundational guide to graph analysis, focusing on graph metrics.

5.1 Introduction

The human brain is one of the most complex systems in the world, consisting of approximately 86 billion neurons. These neurons are interconnected by around 150 trillion synapses and communicate through the transmission of signals, including both electrical and chemical signals, in order to generate human behaviours [14][153]. Understanding the underlying mechanism of human cognition and behaviour under various conditions, including normal and abnormal states, is an extremely important goal of neuroscience. Recently, numerous studies have considered modeling and analysing the brain from the perspective of connectivity patterns. This approach provides comprehensive information behind human cognition, behavior, and perception [155, 137, 174]. Brain connectivity patterns can be divided into the structural and functional connectivity. However, the majority of studies have been devoted to functional connectivity [31].

There are many tools for analyzing connectivity patterns; however, due to the complexity of brain connectivity patterns, as well as its network nature, graph theory is more commonly used. This theory has become the most popular tool for analyzing brain connectivity patterns [198, 201]. Graph theory provides a robust mathematical framework to model and investigate brain connectivity from the network point of view. The graph-based network analysis has the potential to reveal meaningful information about the topological architectures of the human brain and their changes under various conditions, including normal and abnormal states [35, 202]. These changes can provide insight into the underlying mechanism of human cognition, and can be used to guide diagnosis and treatments.

Thanks to advancements of neuron imaging, nowadays, neuronal imaging modalities such as diffusion tensor imaging (DTI), functional magnetic resonance imaging (fMRI), Fluorodeoxyglucose Positron Emission Tomography (FDG-PET) can be used to model brain networks and study different disease pathways, contributing significantly to our knowledge about the brain and brain disorders. Among those imaging modalities, fMRI is a most common one and there have been numerous studies using fMRI to model brain networks [224, 66], just to name a few. Compared to fMRI and other neuroimaging modalities, FDG-PET images provide unique information which cannot be achieved by the others. It provides information about glucose consumption in the brain, which is directly correlated with brain synaptic activity [13]. Moreover, metabolic brain networks highlighted by FDG-PET images are not completely explained by functional or structural brain networks [208, 122], even highly correlated. This suggests that metabolic connectivity networks might exhibit distinct properties from functional connectivity, highlighting the importance of brain metabolic connectivity analysis using FDG-PET images. However, the number of studies on metabolic networks and graph theory is still limited compared to fMRI imaging. Moreover, there is a lack of survey articles discussing FDG-PET-based brain connectivity networks and graph theory as tools for analyzing metabolic connectivity patterns.

In this study, we perform a systematic search of literature to review recent studies using graph theory for studying FDG-PET-based connectivity patterns. With this study,

we expect to see whether **FDG-PET-based connectivity and graph theory can help us better understand the underlying mechanism under various conditions, and to provide a short overview about the most common techniques used to construct and analyse metabolic-based networks**. The remaining sections of this article are organized as follows. The Methodology section outlines the criteria for selecting and reviewing papers in this study. In the Graph Background section, we begin with a short description of graph theory (subsection 3.1), summarize established methods to build and explore brain network connectivity (subsections 3.2 and 3.3), then delve into the most common graph metrics to analyse the architecture of brain connectivity (subsection 3.4). The Results section represents findings from the literature search and an overall summary of the chosen articles. Moving on to the Discussion section, we explore the potential implications and applications of graph theory in healthy groups (subsection 5.1) and its relevance to common neurological conditions (subsection 5.2). Finally, the Challenges and Future Directions section addresses current challenges, as well as potential future developments in this rapidly advancing field.

5.2 Methodology

5.2.1 Research Questions

Based on the objectives of this study which is described before in the abstract, the following research questions were derived and form as the basis for this study:

- RQ1: How have methods for modeling brain metabolic patterns with FDG-PET changed over time? ***(Evolution)***
- RQ2: How can we categorize research on mapping the human metabolic connectome using FDG-PET? ***(Categorization)***
- RQ3: How has the application of graph theory concepts evolved in understanding brain metabolism through FDG-PET imaging? ***(Graph Foundations)***
- RQ4: What are the latest advancements in utilizing graph-based approaches to explore metabolic networks using FDG-PET images? ***(Applications)***
- RQ5: What challenges and promising future directions emerge from the current state of graph-based research on metabolic networks with FDG-PET? ***(Challenges and Futures)***

For those RQs, while RQ1, 2 and 3 provide the answers for the important graph theory for FDG-PET-based metabolic network analysis and related knowledge. The RQ4 and 5 give a more detailed picture about related research studies, as well as future directions.

5.2.2 Search Strategy and Eligibility Criteria

The search strategy involves several steps, including determining the search space and exploring relevant materials. In this study, we focus on the studies regarding graph-based methods for analyzing FDG-PET-based metabolic network, including peer-reviewed journal articles, proceedings, conferences articles. The research process is performed on Web of Science with two data collections, including Web of Science Core Collection and MEDLINE. Since initial work with PET that started in the 1980s [165][45], we collect all articles which have been published after 1980, with the following keyword combinations in the title, keywords or abstract: **("FDG-PET" OR "FDG PET") AND ("metabolic network" AND "metabolic graph" OR "metabolic connectomics" OR "metabolic connectivity")**. After searching process, these included articles are ones which satisfied the following criteria: (a) written in English; (b) peer reviewed; (c) identify and describe methods to build, quantify or compare metabolic connectivity patterns in the human brain network; (d) applied to FDG-PET data. Based on searching articles, we perform analysis and summarize results in order to find out the answers for research questions.

5.3 Basic Concepts

5.3.1 Graph Theory

The roots of graph theory and network analysis can be traced back to 18th century when Leonhard Euler solved the Königsberg Bridge Problem. In Euler's work, he described graph as a finite set of nodes that are connected together by edges. Since then, graph theory has been used and applied in a variety of fields, including physics, biology, chemistry, information systems, just to name a few. However, the application of graph theory for brain networks, only became widespread after 2005, following the introduction of the "Human Connectome" by Olaf Sporns and his colleagues [198][201]. From this time point, the term "Brain Connectome" or "Brain Connectomics" became more prevalent.

When working with brain networks, two fundamental questions need to be answered: *(i) methodology for graph construction* and *(ii) the techniques for graph analysis*. Specifically, in the construction phase, there are two important things that need to be taken into account: *(i.1) how do we define the nodes*, and *(i.2) how do we determine the connections between them*. Meanwhile, when dealing with graph analysis, several graph metrics can be computed to reveal graph topologies and characteristics. Studying the changes in graph topologies is expected to provide insights about the brain structure and underlying mechanisms for behaviors [137][35]. In the next paragraphs, we will briefly describe several important properties of our brain from graph theory perspectives

To survive in a changing world, our brains have to process a lot of information and convert it into clear thoughts to help us make smart decisions. This process is related to two important properties of brain networks: *integration* and *segregation* [199].

Integration refers to the ability of different brain regions to work together, allowing efficient exchange of information and the coordination of various cognitive functions. Segregation, on the other hand, is the capacity of the brain to separate and process information within specialized, local networks or modules. Maintaining a balance between integration and segregation is crucial for the optimal functioning of the brain [199, 129].

Another important property of brain networks is the "small-world" property. This concept originated from the work of Watts and Strogatz in 1998 [231], which revealed that various networks in social and biological domains exhibit a unique organization. Small-world networks display characteristics of both regular and random networks, featuring a short average path length between nodes and a high clustering coefficient. This architecture efficiently balances the minimization of resource costs and the maximization of information flow within the network components. The anticipated small-world architecture of the human brain network is attributed to lower metabolic and wiring costs in connections among nearby brain areas compared to distant regions. Both theoretical and empirical studies support the notion that brain regions tend to interact more with their neighbors to reduce overall metabolic costs, while maintaining a small number of long-distance connections to enhance the efficiency of data transmission [16, 138, 34].

In addition to integration, segregation, and small-world properties, there are other graph metrics employed to assess the topological properties of networks, such as shortest path length, global coefficient, hub, and so on [180]. They all play a crucial role in network-based studies. Each of these metrics has its own advantages and disadvantages. While it's challenging to determine the most suitable measures for studying the complex structure of the human brain, those capable of representing small-world properties and hub topologies are deemed significant. In the following subsections, we will briefly describe the process of building a brain connectivity network using FDG-PET data, as well as the most common graph measures found in the literature.

5.3.2 Construction of Metabolic Connectivity network using FDG-PET

After preprocessing FDG-PET images, before further analysis, we need to construct a graph. This procedure involves several steps. Firstly, we need to determine a suitable parcellation scheme such as using brain atlas for dividing the entire brain into smaller parts, known as region of interests (ROIs). The choices of atlas can be varied, however, Automated anatomical labelling (AAL) atlas [216] can be considered as the most common one. The ROIs, then, can be considered as the nodes in the graph. After that, we need to compute the edges/links between nodes. This step can be done by using correlation metrics such as: covariance, Pearson scores and so on. After these steps, we obtain a correlation matrix which illustrates relationships between vertices in a network. Brain graphs can be either directed or undirected, depending on the presence of directional information in the links. Furthermore, brain graphs can be

weighted or binary, with weighted networks incorporating diverse information, such as fiber number and length, obtained from techniques like diffusion MRI. In order to facilitate further analysis, the correlation matrix is usually converted into a binary connectivity matrix using a specific threshold value (often values to keep 10 or 20 percent of strongest connections). Finally, several key topological metrics of graph can be computed and used to compare graphs together.

5.3.3 Computing Graph Measures

In graph theory, an adjacency matrix illustrates relationships between vertices in a network. Extracting metrics from this matrix enables the topological analysis of graphs, such as the human brain network. Due to technical constraints, most human brain investigations have focused on undirected networks. Graph metrics can be divided into two groups: global and local metrics. The differences between the two groups are: while global metrics reveal global characteristics of the network such as segregation, integration, small-worldness, network resilience; local metrics give information about most significant nodes which have a lot of contribution to network topology.

Global metrics Global measures in brain network analysis aim to uncover several key aspects: (a) functional segregation, which assesses how network elements form specialized communities; (b) functional integration, providing insights into the efficiency of global information communication; (c) small-worldness, indicating an optimal balance between network segregation and integration; and (d) network resilience against failure. Segregation is measured by clustering coefficients and modularity, which quantify topological properties in brain networks, considering anatomically adjacent or functionally connected areas as modules. Various studies show that networks based on modular structures often reflect small-world properties. Functional integration is typically assessed by the characteristic path length, representing the ability for global information integration. The small-world property ensures an optimal balance between segregation and integration, facilitating efficient communication across the network. Additionally, assortativity measures the network's resilience against random or deliberate damages in its main components, addressing a significant concern in network science.

Local metrics In network science, hubs are nodes with high nodal centrality, significantly influencing the network topology. There are two types of hub nodes: connector hubs, which interconnect nodes between different modules, and provincial hubs, responsible for linking nodes within the same module. Detecting hubs in a network is commonly done by calculating nodal degree, i.e., counting the edges connected to each node. The degree distribution of a network, represented by $P(k)$, provides insights into the presence of hubs; for example, in scale-free networks, the existence of several high-degree nodes is associated with a power-law distribution.

Other frequently used measures for nodal centrality include betweenness, closeness, eigenvector, participation coefficient, and PageRank. These measures help assess the importance and influence of nodes within a network.

5.3.4 Most common graph metrics

In this section we provide a short introduction to the most popular graph metrics, more details about those metrics can be found in [72]. Formally, a graph can be represented by $G = (V, E)$ with n nodes, $V = \{v_i | i \in 1, 2, ..., n\}$ is the set of all nodes (or vertexes) in the network, E is set of links (or connections, edges) between two nodes $E = \{(v_i, v_j) | i, j \in 1, 2, ..., n\}$, $a_{i,j}$ is the connection status between node v_i and node v_j. $|V|$, $|E|$ denotes the number of nodes and the number edges, respectively. Graph can be classified as directed or undirected based on whether the edges have sense of direction information. Graph can also be divided into unweighted (binary) or weighted graph. If every edges in the graph has an equal value 1, it can be considered as binary graph, while edges are signed with different values, we call it as weighted graph. Since weighted and directed graph can be very complex for visualizing and analysis, most of studies focus on undirected and unweighted graph. In this section we will mainly focus on graph metrics for undirected and unweighted graph. However, Graph metrics for directed and weighted graph can be found in [180] [72].

For an undirected and unweighted graph $G(V, E)$ the connectivity pattern can be described by a symmetric matrix with size $n \times n$, named adjacency matrix A whose element a_{ij} is 1 if there exists an edge between two node i and j and 0 for the rest. Based on adjacency matrix A, now, we will describe the most commonly graph metrics can be used for metabolic connectivity analysis.

5.3.4.1 Network Degree and Network Density

In a graph G, the degree of node i is the number of edges linked to it and is calculated as:

$$k(i) = \sum_{j \in V} a_{i,j}$$

Degree is one of the most basic graph metrics which measuring the connectivity of a node with the other ones in a network. The mean of degrees over all the nodes in G, called average degree, measures the extent to which the graph is connected. The higher value of average degree indicates the more connected network. Network density, or network density score, is another simple metric that quantifies the proportion of realized connections among all possible connections in a network:

$$D = \frac{2 \times |E|}{n \times (n-1)}$$

It provides insights into how well-connected or sparse a network is. The network density takes a value in range from 0 (no edges) to 1 (fully connected graph).

5.3.4.2 Shortest path length and Characteristic Path Length

Shortest path length, $d_{(i,j)}$, between two nodes i and j measures the minimum number of edges needed to traverse between two nodes. It quantifies the efficiency of information transfer between two nodes. It is computed as

$$d(i,j) = \sum_{a_{u,v} \in g_{i \longleftrightarrow j}} a_{u,v}$$

where $g_{i \longleftrightarrow j}$ is the shortest path between node i and j. The smaller value of shortest path length indicates a more efficient and well-connected network. Or in the other words, nodes can be reached more quickly with very few steps. Based on shortest path length, the characteristic path length (L) of a network is defined as the average shortest path length between all pairs of nodes in the graph:

$$L = \frac{1}{n(n-1)} \sum_{i \neq j} d(i,j)$$

It is a global graph metric indicating how well information can be exchanged between different parts of the network.

5.3.4.3 Network efficiency

Efficiency is a more biologically relevant metric to describe brain networks from the perspective of information flow. It quantifies how efficiently information can be exchanged between nodes in the network. The global efficiency (E) is defined as average of the inverses of the shortest path lengths between all pairs of nodes:

$$E = \frac{1}{n(n-1)} \sum_{i \neq j} \frac{1}{d(i,j)}$$

The local efficiency of G is measured as:

$$E_{loc} = \frac{1}{n} \sum_{i \in G} E(G_i)$$

where $E(G_i)$ is the global efficiency of sub-graph G_i which contains the neighbors of node i. Similar to characteristic path length, global and local efficiency are global metrics measuring how efficiently information can be transmitted between nodes. A higher efficiency value indicates a more efficient network for information exchange.

5.3.4.4 Number of Triangles and Clustering coefficient

The number of triangles around a node i (t_i) captures the presence of interconnected triads. It signifies a strong clustering pattern and reveals the extent to which vertices

tend to form interconnected groups:

$$t_i = \frac{1}{2} \sum_{j,h \in V} a_{i,j} \times a_{i,h} \times a_{j,h}$$

Based on the number of triangles, the clustering coefficient of node i (C_i) can be defined as:

$$C_i = \frac{2 \times t_i}{k_i \times (k_i - 1)}$$

and clustering coefficient of network G equals to:

$$C = \frac{1}{n} \sum_{i \in V} C_i$$

5.3.4.5 Nodal centrality

Nodal centrality quantifies how important a node is within a network. A node with high centrality is considered a hub in the network which means that it is a highly connected node, serving as a central point in the network, facilitating efficient communication and information transfer between different parts of the network. The removal or changes of a hub can have a large impact on network connectivity. In order to measure nodal centrality, several metrics can be used such as nodal degree, closeness centrality, and betweenness centrality. In which, the closeness centrality of node i is computed as

$$c_i = \frac{n-1}{\sum_{j \in V, j \neq i} d_{ij}}$$

and the betweenness centrality of node i is equivalent to

$$b_i = \frac{1}{(n-1)(n-2)} \sum_{\substack{m,n \in V \\ m,n \neq i}} \frac{\sigma_{mn}(i)}{\sigma_{m,n}}$$

, where σ_{mn} and $\sigma_{mn}(i)$ are the number of shortest paths between node m to node n, the total number of shortest paths from node m to node n that pass through node i, respectively. While the closeness centrality reflects the average distance from a node to all other nodes in a network, the betweenness centrality captures the influence that one node has over the flow of information between all other nodes in the network.

5.3.4.6 Small-world

A small-world network is a type of network that exhibits two key characteristics: high clustering and short average path lengths. In other words, in small-world networks, most nodes are closely connected to their neighbors, which promotes local information processing and efficient communication within clusters. At the same time, there are a few "long-range" connections that bridge different clusters, facilitating global communication and information transfer across the network.

Small-worldness can be quantified by comparing a given network's clustering and average path lengths to those of a random network with the same number of nodes and edges. The most commonly used measure to quantify small-worldness is the small-world index (σ), which is defined as the ratio of the clustering coefficient (C) to the average path length (L) of the network divided by the clustering coefficient (C_{rand}) and the average path length (L_{rand}) of a random network with the same number of nodes and edges:

$$\sigma = \frac{C/L}{C_{rand}/L_{rand}}$$

or

$$\sigma = \frac{\gamma}{\lambda}$$

in which $\lambda = \frac{L}{L_{rand}}$, and $\gamma = \frac{C}{C_{rand}}$ If the small-world index (σ) is greater than 1, the network is considered to have small-world properties. A value of σ much greater than 1, which corresponding to $\gamma \gg 1$ and $\lambda \sim 1$, indicates that the network has a high degree of clustering and short average path lengths compared to a random network, which are characteristic features of small-world networks.

This section is partly addressed the research question 3. More details regarding the construction of brain networks, graph metrics and way to interpret their meanings can be found in [180, 72, 31].

5.4 Results

5.4.1 Literature Search

In order to select articles for this review, we followed steps from the PRISMA guideline [142] including identification, screening and selection. In the first step, 68 articles were identified. The oldest article is from 1998 and the number of articles published before 2008 is 5, accounting for only 7 percent of all papers. Most of the articles are published during the period from 2014 to the present, accounting for approximately 84 percent, reflecting the growing interest in the research of brain metabolic patterns and FDG-PET images. In the second step, relevant articles were selected from the 68 articles using an abstract screening process with inclusion criteria, as mentioned in section 5.2. We mainly focus on neuroimaging and graph theory with applications in the human brain as research subject, removing any research regarding animals, as well as non-graph-based approaches. After these steps, 28 articles were selected for full-text verification and detailed review. Details about these selected articles can be found in Table 5.1. Figures 5.1, 5.2 describe the publication trend and categorization based on data group, respectively.

5.4.2 Study Characteristics

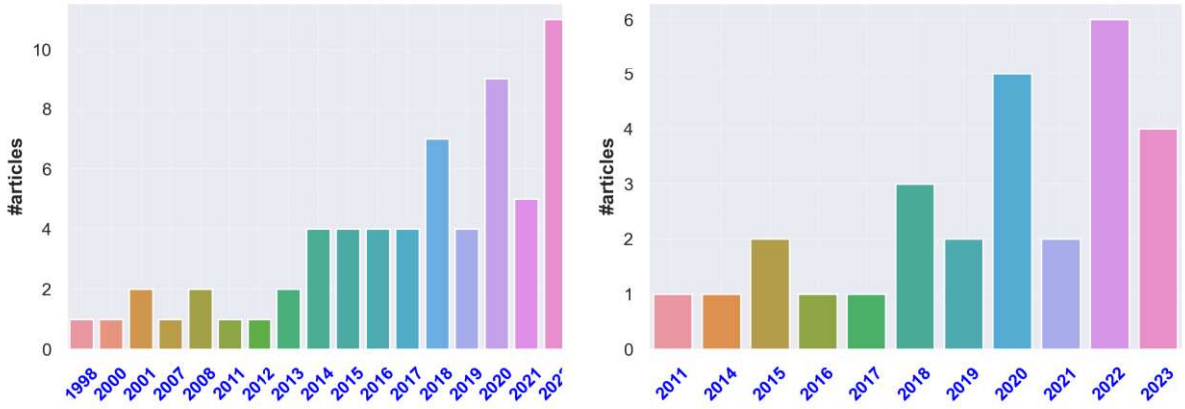

Figure 5.1: Publication trend by year of keyword-matched studies (left) and included studies (right).

The sample size across studies varied from 29 to 700 participants. The mean and standard deviation of participants in all study samples were 206 and 172, respectively. Depending on the selected criteria, these studies can be divided into groups based on the type of methods for constructing graphs (group-based graph, individual-based graph) or patient conditions (Normal condition or Abnormal conditions). More specifically, from FDG-PET images, previous studies either constructed a single graph for each group of data or an individual graph for every image. Different techniques can be applied, subsequently creating different types of group-based graphs [118, 99, 103], and individual-based graph networks [238, 101, 227]. For patient conditions, the studies can be performed with only a healthy control group. These types of studies typically focus on studying the correlation between FDG-PET network-based data and metadata such as subject's gender [94], age [97, 98]. They may also investigate the correlation and similarity between FDG-PET-based network and network from other modalities such as sMRI [122]. When patient group data is available, most of these studies focus on either investigating the topology changes in brain networks [99, 101] or using connectivity and machine learning for classification tasks [238, 124, 123]. In the next sections, we will go further into details and discuss these studies.

5.4.3 General Overview

In this section, we present an overview about the selected articles, focusing on publication year and keywords analysis. This discussion addresses the research question 1, 2, and 3.

From Fig. 5.1 which illustrates the growing interest in research topics, we can observe that there is an increasing trend. Among the articles matched by keywords,

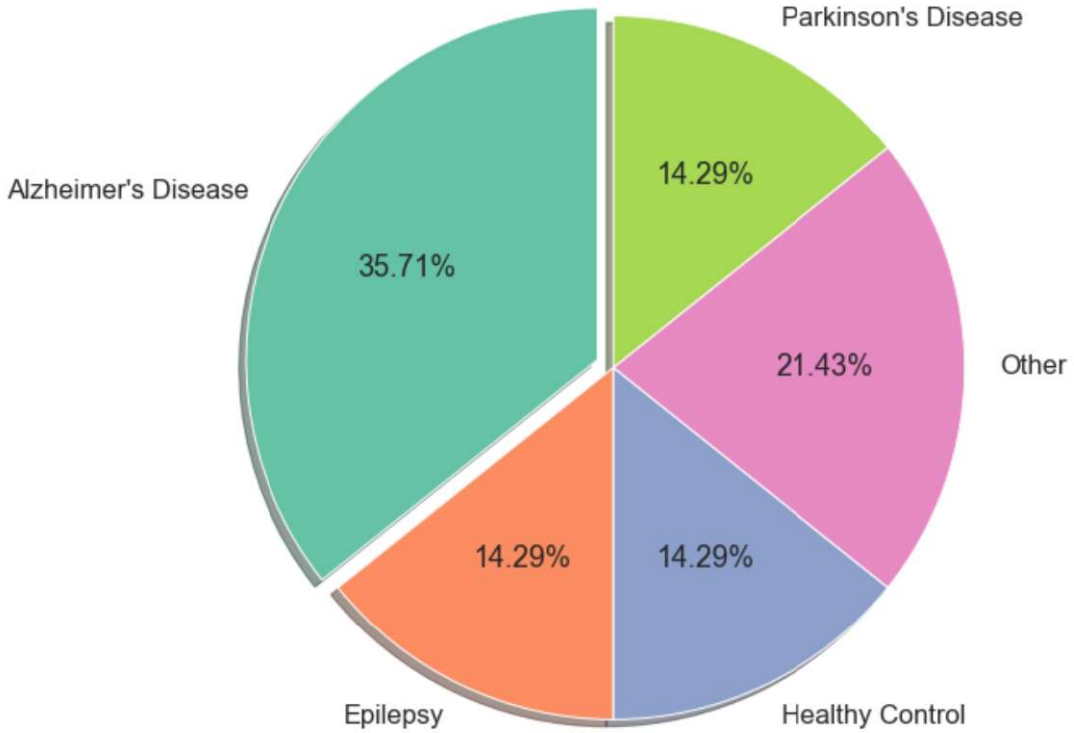

Figure 5.2: Categorization of included studies based on data group.

most studies belong to the period from 2014 to the present. Simultaneously, among the selected articles, most of them belong to the period from 2018 to the present. Comparing the periods from 2011 to 2017, the number of publications from 2018 to the present has increased approximately fourfold. This trend is expected to show a substantial increase in the coming years as FDG-PET becomes more accessible.

Fig. 5.2 categorizes the included studies into 5 groups based on the data they used. It provides a visual breakdown of the prevalence of neurological disorder analysis using graph theory and FDG-PET images. Alzheimer's Disease (AD) emerges as the dominant category, reflecting its widespread occurrence and significant impact on neurological research. Following closely are Parkinson's Disease (PD) and Epilepsy, highlighting their prominence and relevance in the metabolic connectivity studies using graph theory. This observation is reasonable as these disorders are common and of substantial importance to the world healthcare system. AD and other related disorders have attracted a lot of attention from the research community. Moreover, the dominance of AD in our collection can be explained by the contribution and the availability of the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset [163].

Fig. 5.3 and 5.4 present the word-cloud for the top 50 most frequently used words

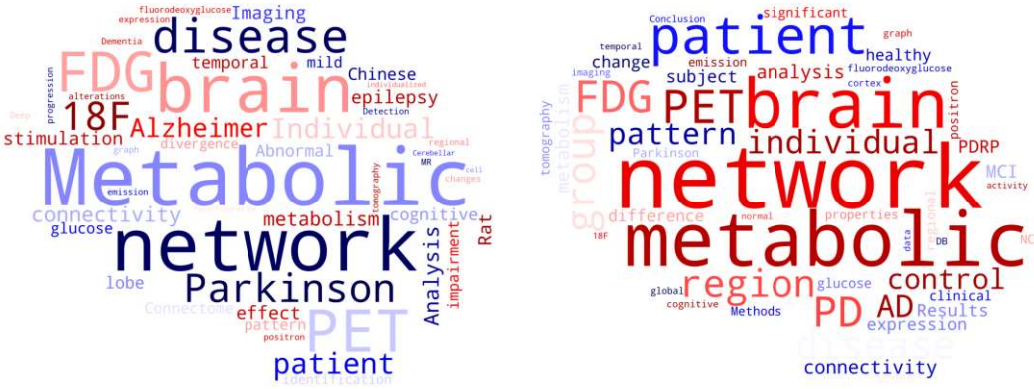

Figure 5.3: Word Clouds illustrate the top 50 of most frequency words in title (left) and abstract (right) of keyword-matched articles.

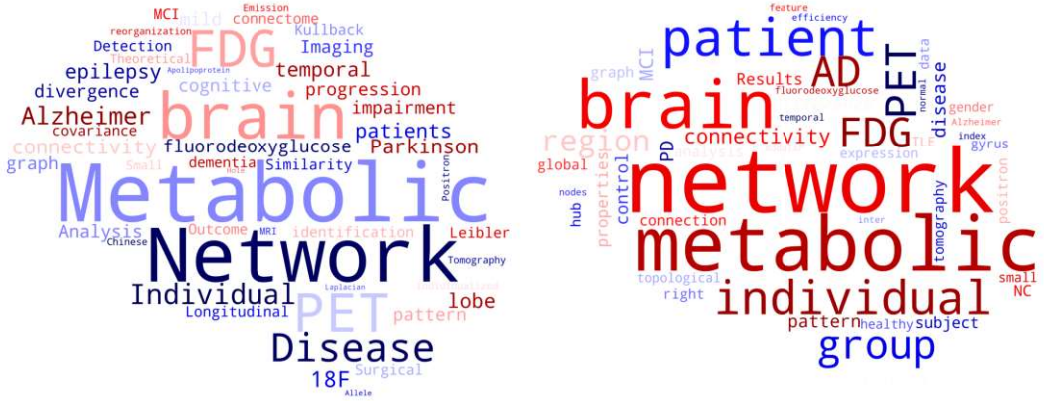

Figure 5.4: Word Clouds illustrate the top 50 of most frequency words in title (left) and abstract (right) of included articles.

from title (left) and abstract(right), corresponding to keyword-matched articles and selected articles, respectively. From Fig. 5.3, we observe that keywords such as "network", "FDG", "PET", "metabolic", "PD", "AD", "epilepsy", are predominant. The first four keywords are reasonable as they were our used keywords for searching relevant articles. The rest are related to most common group applications, which is in line with observation of the prevalence of neurological disorder analysis using graph theory and FDG PET images from Fig. 5.2.

Comparing Fig. 5.3 to Fig. 5.4, even though they are very similar, we observe some differences. When comparing the top words in the title, we can see that the word "rat" disappears in Fig. 5.4. This observation is reasonable since we have removed a lot of studies related to rat models, as those studies are not related to the human brain. When comparing the top words in abstract, we can observe changes, including: (i) The appearance of new words such as topological, efficiency, and hub. Interestingly, they are all related to graph and graph metrics. This observation emphasizes the important

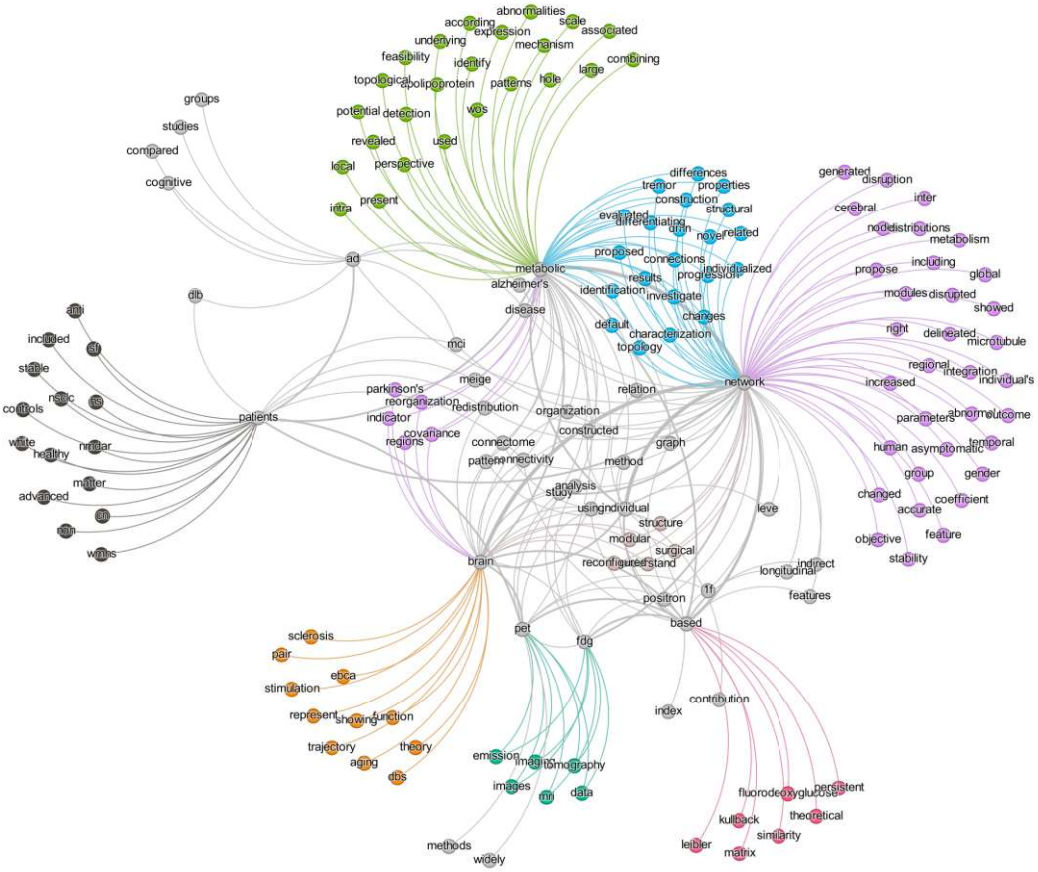

Figure 5.5: Collocation Analysis for words in Abstracts of included articles.

role of topological changes in selected studies. (ii) The keywords "individual" becomes more dominant in Fig. 5.4. This suggests a growing interest in applying graph theory to analyze differences between each individual under different conditions.

Fig. 5.5 depicts the collocation analysis illustrating the correlation between words in the abstracts of selected papers. It provides us with information regarding the relationships between words and their co-occurrence patterns within the given context. It is notable that certain keywords such as "metabolic," "brain," and "network" emerge as the most salient collocates. This observation aligns with findings from Figs. 3 and 4, which is consistent given our studies' primary focus on the metabolic brain network. Additionally, keywords like "patient" and "ad" appear more frequently. Given that most studies concentrate on Alzheimer's disease, it is reasonable to encounter "ad" multiple times. Regarding "patient," its frequent appearance can be attributed to syntactic constraints, as most studies involve non-healthy groups. Further examination of this figure can reveal additional insights. For instance, the most common keywords associated with "network" and "metabolism" include "changes," "differences," "topology," "connections," and "properties," suggesting that many studies focus on analyzing network properties and topology to identify differences or changes

between groups.

5.5 Discussion

Given that FDG-PET offers insights into brain glucose consumption, which is directly correlated with brain synaptic activity [13], quantifying the degree of association between pairs of brain regions (referred to as metabolic connectivity), is expected to provide early indications of abnormal brain states. The analysis of metabolic connectivity networks using graph theory can be applied to two groups i) Healthy controls; ii) Neurological disorders. The first group mainly focuses on investigating the neural mechanism that underlie human cognitive function and behaviour through the lens of brain ROIs interactions. While the second group focuses on assumptions about disconnection phenomena found in brain disorders. In the next two subsections, we will discuss more details about these two groups. This section aims to figure out the answers to research question 4 and 5.

5.5.1 Healthy Control-based Metabolic Network Analysis

In exploring brain connectivity within the healthy control group, previous studies reveal valuable insights into the intricate organization of human brain networks. More specifically, in the study [122], authors concentrate on cross-modality similarity and dissimilarity between metabolic and structural brain networks. Based on comparing graph metrics computing from metabolic and structural graphs, it shows that both networks exhibited small-world architecture; highlight the central role of default mode network. The study came to the conclusion that there was a structural substrate that supported the brain to globally and efficiently integrate and process metabolic interaction across brain regions. The cross-modality cooperation or specialization in both metabolic and structure networks might imply mechanisms of achieving higher-order brain functions.

Other studies [94, 97, 98] have considered the effects of age and gender to metabolic brain organization. Study [98] delves into aging-related reorganization, noting shifts in modular architecture and reduced connector nodes in older age, particularly in temporo-occipital areas. Study [97] focuses on brain developmental trajectories. Based on the analysis of metabolic networks from hundred subjects across different age groups, the study shows that both network connectivity and global efficiency increase with age. Brain networks were grouped into modules, showing that modular architecture was dramatically reorganized from childhood to early adolescence. Hubs spatiotemporally rewired, and the ratio of connector hub to provincial hub also changed with age. Based on these studies [97, 98], it appears that topological properties and modular reorganization of human brain networks dramatically change with age, especially from childhood to early adolescence. Not only age but gender might also modify brain metabolic connectivity. Study [94] explores gender differences in brain metabolic networks. Based on the analysis of 400 healthy subjects, study [94]

showed while both genders exhibited small-world properties in metabolic networks with no significant global gender differences, nodal differences were observed in a few brain regions. Hubs were distributed bilaterally and showed lateralized patterns in females and males.

In comparing those studies together, all studies leverage graph theory to explore the topological properties and modular architectures of brain networks, emphasizing the applicability of this analytical framework for understanding complex neural systems. Moreover, effects of age and gender on connectivity highlighted in the above studies suggest the need for a multifaceted approach when studying brain connectivity. One should consider age, gender, and their interplay in shaping the intricate architecture of the healthy brain.

5.5.2 Metabolic Network Patterns in Brain Disorders

Many brain disorders can be conceptualized as a disconnection phenomenon within brain networks. In [38], authors detail disconnection syndromes, which occur when localized and specialized areas are disconnected to each other, leading to dysfunction. The disconnection phenomenon has been successfully applied to analyze various types of brain disorders. For instance, in [236] and [171], authors reported that based on the analysis of network properties of metabolic brain networks using the graph theory, AD subjects could be distinguished from control subjects. In the following, we will describe several studies which have used graph theory and FDG-PET-based brain networks for the diagnosis of brain disorders.

Epilepsy is a neurological disorder characterized by recurrent, unprovoked seizures. Seizures occur due to abnormal and excessive electrical activity in the brain, leading to temporary disruptions in normal brain function. The manifestations of seizures can vary widely, ranging from brief lapses in attention or muscle jerks to more severe convulsions. Temporal lobe epilepsy (TLE) is a common form of partial epilepsy. Studies [225, 177, 42, 230] on metabolic networks in TLE provide a complementary understanding of altered network dynamics, prognostic indicators, and specific regional changes in this disease. More specifically, study [225] examined the metabolic covariance networks in mesial temporal lobe epilepsy (mTLE). The study highlighted damaging and compensatory effects in metabolic networks of mTLE, - emphasizing the role of reduced inter-hemisphere connection. Study [177] focused on surgical prognosis in drug resistant TLE. This study demonstrates that ongoing postoperative seizures are associated with elevated networks integration, altered connectivity patterns, and redistributed hub nodes. It provides insights into the complexity of TLE cases that do not achieve seizure freedom after surgery. Study [42] investigated differences in metabolic networks among TLE patients with hippocampal sclerosis, based on surgical outcomes following anterior temporal lobectomy: seizure-free and non-seizure-free. Authors found that the seizure free group had a lower assortative coefficient and widespread regional differences. This emphasizes the potential role

of metabolic network analysis on FDG-PET to predict surgical outcomes in TLE patients with hippocampal sclerosis. Study [230] focused on the default mode network (DMN) in TLE. This study reveals abnormal metabolic connectivity patterns in DMN of TLE, particularly in regions like the bilateral posterior cingulate gyrus and hippocampus. It provides insights into the dysfunction mechanism of DMN in TLE as well as pathophysiology of TLE.

The above studies gave interesting insights into the altered network dynamics associated with TLE. They offer distinct perspectives on different aspects of TLE, ranging from surgical prognosis and outcomes to specific network alterations in DMN; emphasizing the importance of metabolic connectivity analysis in epilepsy.

Parkinson's Disease is a neurodegenerative disorder that primarily affects movement. It is characterized by symptoms such as tremors, stiffness, slowness of movement, and difficulty with balance and coordination. The condition results from the gradual loss of dopamine-producing neurons in the brain. Previous studies using FDG-PET-based metabolic network and graph theory analysis have emphasized the important role of FDG-PET images as a valuable tool for both diagnosis and treatment evaluation of PD. Authors in [143] revealed a Parkinson's disease tremor-related metabolic network. When investigating metabolic brain networks of PD patients with tremor, they found a distinct pattern involving cerebello-thalamo-cortical pathways. This specific metabolic network, identified through FDG-PET and analyzed using ordinal trends canonical variate analysis, demonstrated excellent reproducibility and responsiveness to interventions, providing a valuable objective measure for evaluating treatment effects on tremor symptoms. The study [76] explored the potential of a metabolic network derived from PET in monitoring subthalamic nucleus deep brain stimulation (STN DBS) treatment for PD. PD patients with STN DBS were longitudinally followed after FDG-PET scans. It showed that Parkinson's Disease Related Pattern (PDRP) expression was elevated in PD patients, decreased significantly post-STN DBS treatment, and correlated with clinical improvements. Based on that, authors concluded that PDRP could serve as an objective biomarker for monitoring and assessing STN DBS treatment in PD patients. The exploration of the rich-club organization in the metabolic connectome of PD patients uncovered alterations in connections, suggesting potential early markers for PD diagnosis [160]. Lastly, the investigation into individual metabolic connectomes using Jensen-Shannon Divergence Similarity Estimation method demonstrated altered nodal properties in PD [123], highlighting the method's promise for accurate PD diagnosis. In conclusion, these studies collectively underscore the power of FDG-PET and graph theory in providing novel insights into the intricate metabolic networks associated with PD, offering valuable tools for both diagnosis and treatment evaluation.

Alzheimer's Disease is a neurodegenerative disorder affecting memory and cognitive functions, leading to dementia. It is characterized by altered structural and functional connectivity, making it a disconnection syndrome. Mild Cognitive Impair-

ment (MCI) is a condition that represents a noticeable decline in cognitive abilities beyond what is expected with normal aging but is not severe enough to be classified as dementia. MCI can be a precursor to Alzheimer's disease, but not everyone with MCI will progress to AD.

FDG-PET and graph theory have been extensively used to study AD, revealing notable alterations primarily observed in the small-worldness properties of metabolic brain networks in AD subjects [237, 240]. More specifically, upon exploring brain network characteristics, the main differences between AD and cognitive normal (CN) groups includes reduced number of connectivity, global efficiency, clustering coefficient, loss of small-world properties. The key patterns can be found in parieto-occipital, and parieto-occipital connectivity. Graph theory applied to FDG-PET data in those studies has also revealed changes in the inter and intra-hemispheric connectivity patterns at both group-based and individual-based networks [99, 101]. These metabolic network patterns not only serve as reliable biomarkers for diagnostic purposes [238, 99], disease progression assessment [245, 227, 118], but also can be used to investigate the relationship between several factors regarding the risk of AD progression. In [237], authors used graph theory and FDG-PET-based metabolic networks to explore the relationship between APOE $\varepsilon 4$ and the risk of AD progression. Based on the comparison of graph metrics between APOE $\varepsilon 4$ carriers and noncarriers, the author concluded that there was a less optimal organization of metabolic networks in APOE $\varepsilon 4$ carriers, which might explain the risk of developing AD. Loss of small-world properties, topological aberrations, and altered connectivity identified through graph theory differentiate CN from AD, and disease states like stable MCI (sMCI) and progressive MCI (pMCI). They also offer interesting insights into the pathophysiological mechanisms driving cognitive dysfunction [240].

The application of machine learning has played a significant role as a powerful tool for better diagnosing AD. Studies such as [238, 124] and [245] have harnessed the capabilities of machine learning as a diagnostic tool, showcasing the potential of metabolic networks to achieve remarkable accuracy in distinguishing disease groups and predicting clinical progression. The roles of FDG-PET in these studies are pivotal, with metabolic networks emerging as sensitive biomarkers for MCI progression [245], determining persistent features for disease [118, 116], disease expression [227]. More specifically, in study [116], authors focused on employing persistent homology from algebraic topology to investigate alterations in network organization across various stages of AD. They developed a novel network measurement, called Kernel-based integrated persistent feature (KBI), to quantify differences in integrated persistent features curves in AD. Compared to standard graph metrics, this new feature showed a superior sensitivity in revealing network disruptions throughout the progression of AD. Based on that, authors suggested KBI as a potential preclinical AD imaging biomarker. In study [227], authors develop a new method to capture individual pathological connectivity and use this method to study metabolic connectome expression between sMCI and pMCI group. This new method provides detailed individual information, greater stability, and is a better predictor of MCI conversion than other clinical assessments. In order to explore disruptions in the modular structure of brain networks, study [118]

investigated persistent homology to detect higher-order topological invariants, called holes, in AD metabolic networks. The combination of FDG-PET, graph theory, and machine learning not only enhances diagnostic precision but also offers a deeper understanding of the complex and evolving landscape of AD, paving the way for more effective diagnostic and therapeutic strategies.

The landscape of Alzheimer's disease analysis using graph theory has undergone an interesting evolution, with a notable shift from traditional group-based network studies to an individual approach. Studies at a group level utilized Pearson correlation to compute graph connections [99, 237, 240], while recent individual-level studies, such as[227, 101], use diverse methods to construct graphs. This approach distinguishes itself by focusing on the unique profile of patients, moving beyond generalized group-based analyses. More specifically, in [238], authors used cube to define node, and Pearson score for computing correlation matrix for AD classification. In another work from the same authors [245], they used brainnetome atlas to define nodes, and combined rough and precision distance to compute correlation matrix and used them for predicting MCI. Both of those methods showed promising results. In study [101], the authors proposed a new method for constructing an individual connectivity matrix based on adjustments to a correlation matrix derived from a group of HC subjects. This method involves normalizing the effect size of differences in metabolism level between brain regions in an individual PET scan and the average activity observed in the HC group. Through experiments, they showed that the insights obtained from the individual network align with the group-based network using Pearson scores. In a recent work [227], authors proposed a novel method for constructing individual brain networks using Kernel Density Estimation and Kullback-Leibler Divergence Similarity Estimation. These methods offer superior performance in predicting MCI conversion, and are adapted to the study of other brain disorders such as PD [123], or even the brain of patients with cancer [241]. These studies collectively underscore the importance of individualized analyses, allowing for a more nuanced understanding of AD heterogeneity. Moreover, the transition from group to individual-based graph studies not only enhances diagnostic precision but also represents a promising trajectory for tailoring interventions to the unique profiles of AD patients.

Others The applications of FDG-PET-based metabolic networks and graph theory are not limited to the most prevalent neurological disorders like AD and PD; they also offer a valuable insight into the complexities of rare disorders and other neurological conditions. The studies on White Matter Hyperintensities (WMHs) [131] reported the impaired metabolic networks in patients with WMHs. The findings reveal a distinctive redistribution of the brain metabolic network in individuals with WMHs. On the other hand, study [127] investigated metabolic network patterns in presymptomatic microtubule-associated protein tau mutation carriers for genetic frontotemporal dementia. Through metabolic network analysis in asymptomatic carriers, the study identifies reconfigured hubs in brain regions, lost hubs in specific regions, and compensation in task-related regions, enhanced metabolic connectivity in salience and default mode networks. Study [126] focused on Meige Syndrome, a rare movement

disorder. Firstly, authors identified hypometabolism in the globus pallidus and thalamus on FDG-PET, which might be a sign that the basal ganglia-thalamocortical motor circuit is altered. Secondly, authors observed abnormal small-world properties and centrality changes int the same regions found hypometabolic, reinforcing the possible roles of these regions in the pathogenesis of Meige syndrome. Study [95] demonstrated altered metabolic connectivity networks in AntiN-methyl-D-aspartate receptor encephalitis, the most common autoimmune encephalitis. It showed the following: reduced global and local efficiencies, longer characteristic path length, enhanced network functions in specific regions, and poor internal connection integration in the left hemisphere.

Another subject of interest which is not a proper brain disorder, is the study of brain function in patients with cancers, and more specifically, non-small cell lung cancer (NSCLC). In [241], authors found significant differences in the degree and betweenness distributions of brain network nodes between the NSCLC and control groups. Abnormal brain network connections and nodes were identified in the temporal lobe, frontal lobe, and limbic system of patients with advanced NSCLC. The predictive model, constructed using these abnormal brain network features, showed a significant correlation with actual survival times of NSCLC patients. Based on these observations, authors suggested that individual metabolic brain networks could provide additional clinical information to guide further treatment. However, more studies and larger sample sizes are necessary to confirm this suggestion. Nevertheless, it still emphasizes the significance of metabolic networks in understanding cognitive dysfunctions in cancer patients.

Multiple/Cross disorder Analysis Given the intricate landscape of neurological disorders, it's essential to acknowledge that multiple diseases can coexist within an individual, giving rise to complex interplay. Additionally, some brain disorders may exhibit similar symptoms and affect the same brain regions. This realization underscores the critical need for a cross-disorders approach to provide a holistic perspective on the interconnected nature of neurological conditions. Considering this, in study [103], the author investigated the differences between AD patients, dementia with Lewy bodies (DLB) patients, and CN individuals. The main findings are the whole metabolic network was preserved in CN individuals, while AD patients exhibited diffusely decreased connection, and DLB patients showed a partially but more deeply decreased connection. It also revealed significant metabolic topology differences in the right posterior cingulate and the left transverse temporal gyrus between AD and DLB. This study not only sheds light on commonalities and differences in affected brain regions but also emphasizes promise cross-disorder approach to uncovering shared pathophysiological mechanisms, paving the way for more effective interventions different brain disorders.

5.6 Challenges and Future Directions

5.6.1 From Data Perspective

Limited number of related studies: Although FDG-PET demonstrates great potential for analyzing and studying brain disorders, the application of graph theory in analyzing metabolic-based networks is still limited compared to other neuroimaging modalities such as fMRI [66]. Given that FDG-PET can be considered as a proxy for regional neuronal activity [171], it can provide early signs of abnormal brain states. In the future, it would be interesting to explore the effectiveness of graph theory and FDG-PET in other brain and psychiatric disorders.

Small sample size: Many previous studies have had small sample sizes (smaller than 50 subjects for each group of data) [225, 177, 143]. In order to archive more robust observations and generalized results, it is important to increase the dataset size. Additionally, longitudinal studies are also crucial for improved diagnosis and treatment monitoring.

Mono scale: Most previous studies solely rely on static FDG-PET images, implying a single modality and a single time point. However, it would be interesting to create a network that takes integrated information from a single modality but at multiple time points (e.g., dynamic PET images) and scales (e.g., resolutions), or from multiple modalities. Recent studies demonstrate that FDG-PET images combined with sMRI or fMRI can offer complementary views [208, 122]. Therefore, combining networks from multiple modalities is expected to provide a more comprehensive understanding about higher-order brain functions.

Mono nature: Most prior studies have focused on a single disorder at a time. However, recent research on cross-disorders [217][53] indicate that multiple diseases can have similar effects on the same brain regions and being able to distinguish them is extremely important in order to provide effective treatment. Hence, cross-disorder analysis is a crucial research direction. It is necessary for better understanding and distinction between various disorders, ultimately leading to improved diagnosis.

5.6.2 Building and Analyzing brain metabolic networks

Group-based or Individual-based Network Analysis: When building metabolic-based networks from FDG-PET images, previous studies have either followed group-based or individual-based approaches. Individual approaches provide more precise information, contributing to better diagnoses [227] and can be considered a step towards individualized treatments or precision medicine. However, it is challenging to build an individual graph from single static FDG-PET images, as the methods used need to capture both group properties and subject variances. Therefore, for future work, it would be relevant to study and explore optimal methods to build individual graphs. Moreover, it is also important to compare multiple approaches and to compare the significance between group-based approaches and individual methods to understand the similarities and differences between them.

Brain network nodes: It's obvious that the choice of nodes greatly affects the analysis results. In the literature review, the most common choice for defining nodes is the use of predefined brain atlases such as AAL, as in [99][227, 94]. Besides using predefined atlas for defining nodes, the cube-based approach can be used to define the network nodes. This approach can be considered as an intermediate level analysis between voxel-based nodes, and region-based nodes [238]. It can both represent details about nodes and overcome the limitation of using voxel-based approach. Moreover, using adapted brain mapping, which is more adapted to a specific type of brain disorders [212] could be an interesting approach for future study. Lastly, examining the network at various node scales (e.g., ranging from hundreds to thousands of nodes, or possibly more) can provide a complementary view of the disease, revealing subtle patterns and connections that may not be apparent at a single scale.

Brain network connections: Similar to the selection of graph nodes, the choice of connectivity measurements also significantly impacts graph construction and analysis. In previous studies, for group-based analysis, Pearson correlation has been one of the most commonly used connectivity metrics, as found in [116, 94, 97]. Concerning individual-based analysis, most common used metrics are information-based distance measures such as Kullback-Leibler divergence [227, 241] and Jensen Shannon divergence [123, 122]; even though there is still room for exploring alternative correlation methods and similarity measurements. It is also important to assess the effectiveness of these approaches, their stability regarding artifacts, and their ability to perform well across multiple scales.

Furthermore, all the studies discussed so far have primarily focused on the connectivity between two brain regions, often referred to as pairwise connections or second-order relations. However, this approach tends to overlook the interactions among multiple ROIs, known as higher-order relations. Considering that brain networks are inherently more complex and involve intricate relationships, exploring higher-order connections can provide valuable insights into the non-random attributes of these networks [200]. Tools from hypergraph theory [29] can be utilized to represent and study these higher-order relations. For instance, in a recent study [109], authors investigated a hypergraph-based approach to construct a hypergraph from fMRI data and its applications in AD diagnosis. The results of this study demonstrated that a hyperconnectivity-based network not only enhances classification performance but also helps identify crucial disease-related biomarkers. Therefore, it is worthwhile to explore the potential applications of hyperconnectivity and hypergraph theory in the analysis of metabolic networks.

Graph metrics: Since, graph metrics play a pivotal role in extracting meaningful information from the complex network structures, selecting optimal graph metrics is an extremely important and crucial step before the analysis. There is no rules for selecting optimal graph metrics, it depends on the type of study and its objectives. Most commonly used graph metrics are shortest path length, global coefficient, local efficient, small-worldness. Other types of metrics such as motif measurements and resilient measurements can be used depending on the purpose of the study [180]. Integrating results from different studies also raises a big challenge. Graph metrics

can exhibit variability with different network thresholds, across diverse sources of datasets, and in the presence of varied graph construction methods. This variability hampers the reliability and generalizability of findings. Additionally, addressing the need to establish a clear connection between graph metrics and cognitive meaning is crucial. As we delve into the intricate details of brain network analysis, elucidating how specific graph metrics relate to cognitive functions becomes essential for translating the findings into clinically relevant insights.

5.6.3 Neural Network as a tool for exploring brain connectivity

Studying brain disorders across various scales and resolutions is crucial, yet it introduces complexities in understanding networks and determining connectivity patterns. To address this challenge, the application of neural networks becomes paramount. Their inherent ability to autonomously learn pertinent features from data offers a promising solution. Furthermore, within the realm of Deep Learning, a specific subset known as Graph Neural Networks (GNNs) [18] has garnered significant interest from the research community. Hence, employing deep learning and GNNs as tools for analyzing metabolic networks appears particularly relevant. This approach holds the potential to enhance our comprehension of underlying patterns and correlations between different groups, paving the way for more effective analyses and interpretations.

5.6.4 Clinical Intervention

Transition from neuroimaging and graph theory analyses to practical clinical applications is crucial for a better diagnosis and treatment. Translating insights gained from intricate graph analyses into actionable clinical diagnoses and treatment strategies is paramount for the field's advancement. Integrating neuroimaging features into disease classification might help to define novel sub-types classification, based on pathophysiological mechanisms and not only clinical symptoms.

Following diagnosis, the predictive and monitoring capabilities of neuroimaging and graph theory might help to better grasp disease progression, enabling accurate tracking of changes over time or in response to interventions. This not only refines diagnostic precision but also establishes a foundation for tailoring treatments to individual patient needs.

Overcoming these challenges will be pivotal for advancing the utility of FDG-PET and graph theory in unravelling the complexities of neurological disorders and cognitive processes.

5.7 Conclusion

This paper offers a brief overview of graph theory and its application to the analysis of FDG-based metabolic networks. We begin with a concise description of graphs and graph metrics. Following that, we provide a summary of recent work on metabolic

networks and graph theory, covering various aspects such as types of graphs and targeted brain disorders. And in the final section, we provide a brief description of several challenges and emphasize some interesting research directions for the future.

5.8 Chapter Conclusion

This chapter has provided a brief overview of graphs and graph analysis. In the subsequent chapter, we will delve into the applications of graph theory for analyzing individual brain networks, specifically focusing on its relevance in AD diagnosis.

Table 5.1: List of included studies in this survey

	Title	Year	Subject Group	Data Size	Data source	Ref
1	Hole Detection in Metabolic Connectivity of Alzheimer's Disease Using k-Laplacian	2014	Alzheimer's Disease	91	Not specify	[118]
2	A FDG-PET Study of Metabolic Networks in Apolipoprotein E ε4 Allele Carriers	2015	Alzheimer's Disease	330	ADNI	[237]
3	Individual Metabolic Network for the Accurate Detection of Alzheimer's Disease Based on FDG-PET imaging	2016	Alzheimer's Disease	222	ADNI	[238]
4	Predicting MCI Progression with Individual Metabolic Network Based on Longitudinal FDG-PET	2017	Alzheimer's Disease	79	ADNI	[245]
5	Indirect relation based individual metabolic network for identification of mild cognitive impairment	2018	Alzheimer's Disease	365	ADNI	[124]
6	Characteristic patterns of inter- and intra-hemispheric metabolic connectivity in patients with stable and progressive mild cognitive impairment and Alzheimer's disease	2018	Alzheimer's Disease	300	ADNI	[99]
7	Learning Metabolic Brain Networks in MCI and AD by Robustness and Leave-One-Out Analysis: An FDG-PET Study	2018	Alzheimer's Disease	272	ADNI	[240]
8	Metabolic Brain Network Analysis of FDG-PET in Alzheimer's Disease Using Kernel-Based Persistent Features	2019	Alzheimer's Disease	700	ADNI	[116]
9	A Novel Individual Metabolic Brain Network for 18F-FDG PET Imaging	2020	Alzheimer's Disease	300	ADNI	[101]
10	Individual brain metabolic connectome indicator based on Kullback-Leibler Divergence Similarity Estimation predicts progression from mild cognitive impairment to Alzheimer's dementia	2020	Alzheimer's Disease	610	ADNI	[227]
11	Metabolic covariance networks combining graph theory measuring aberrant topological patterns in mesial temporal lobe epilepsy	2019	Epilepsy	41	Local	[225]
12	Metabolic Brain Network and Surgical Outcome in Temporal Lobe Epilepsy: A Graph Theoretical Study Based on 18F-fluorodeoxyglucose PET	2021	Epilepsy	79	Local	[177]
13	Metabolic network is related to surgical outcome in temporal lobe epilepsy with hippocampal sclerosis: A brain FDG-PET study	2022	Epilepsy	75	Local	[42]
14	Abnormal metabolic connectivity in default mode network of right temporal lobe epilepsy	2023	Epilepsy	81	Local	[230]
15	Small-worldness and gender differences of large scale brain metabolic covariance networks in young adults: a FDG PET study of 400 subjects	2015	Healthy Control	400	Local	[94]
16	Age-associated reorganization of metabolic brain connectivity in Chinese children	2020	Healthy Control	201	Local	[97]
17	Aging-Related Modular Architectural Reorganization of the Metabolic Brain Network	2022	Healthy Control	432	Local	[98]
18	Cross-modality comparison between structural and metabolic networks in individual brain based on the Jensen-Shannon divergence method: a healthy Chinese population study	2023	Healthy Control	209	Local	[122]
19	Metabolic Network Topology of Alzheimer's Disease and Dementia with Lewy Bodies Generated Using Fluorodeoxyglucose Positron Emission Tomography	2020	Other	205	Local	[103]
20	Regional metabolic and network changes in Meige syndrome	2021	Other	100	Local	[126]
21	Brain Metabolic Network Redistribution in Patients with White Matter Hyperintensities on MRI Analyzed with an Individualized Index Derived from 18F-FDG-PET/MRI	2022	Other	120	Local	[131]
22	Cerebral Metabolic Network in Patients With Anti-N-Methyl-D-Aspartate Receptor Encephalitis on 18F-FDG PET Imaging	2022	Other	34	Local	[95]
23	Reconfigured metabolism brain network in asymptomatic microtubule-associated protein tau mutation carriers: a graph theoretical analysis	2022	Other	83	Local	[127]
24	Construction of an individualized brain metabolic network in patients with advanced non-small cell lung cancer by the Kullback-Leibler divergence-based similarity method: A study based on 18F-fluorodeoxyglucose positron emission tomography	2023	Other	138	Local	[241]
25	Parkinson's disease tremor-related metabolic network: Characterization, progression, and treatment effects	2011	Parkinson's Disease	29	Local	[143]
26	Metabolic network as an objective biomarker in monitoring deep brain stimulation for Parkinson's disease: a longitudinal study	2020	Parkinson's Disease	84	Local	[76]
27	Alternation of the Rich-Club Organization of Individual Brain Metabolic Networks in Parkinson's Disease	2022	Parkinson's Disease	98	Local	[160]
28	The reconfiguration pattern of individual brain metabolic connectome for Parkinson's disease identification	2023	Parkinson's Disease	98	Local	[123]

Chapter 6

Metabolic Individual Brain Network Analysis with Graph theory

Summary

6.1 Introduction . . . 129
6.2 Materials and Methods . . . 130
6.2.1 Dataset . . . 130
6.2.2 Methods . . . 131
6.2.2.1 Individual brain network . . . 131
6.2.2.2 Graph Metrics & Graph Comparison . . . 132
6.3 Results and Discussion . . . 133
6.3.1 Results . . . 133
6.3.2 Discussion . . . 134
6.4 Conclusions . . . 136
6.5 Chapter Conclusion . . . 136

In recent times, there has been a surge in studies concentrating on modeling and analyzing the brain's connectivity patterns using FDG-PET images. However, most of these studies have emphasized group-based network analysis. While the group-based approach has yielded several interesting results, methods focused on individual networks offer more precise information, thereby enhancing diagnostic accuracy and tailoring treatment strategies to individual needs. Moreover, compared to other neuroimaging modalities, the utilization of FDG-PET images and graph theory remains relatively limited, indicating a potential avenue for further exploration and research. This chapter aims to explore the contribution of graph theory in analyzing metabolic patterns using FDG-PET images of subjects under various conditions, with a specific emphasis on individual-level analysis. This work is expected to lay the groundwork for individualized analysis and treatment. Specifically, in this thesis, we developed a method for constructing individual brain networks from FDG-PET images, primarily focusing on Alzheimer's Disease (AD) diagnosis.

An effective method for constructing individual graphs should capture the essential characteristics of the group while accurately representing individual variances. In the initial phase of this thesis, we introduced a method for building individual brain

network, named Dynamic Time Warping (DTW)-based distance, a well-established method for comparing two time series signals, to construct individual metabolic networks. Experimental results demonstrated that our proposed methods can effectively capture both group properties and better represent individual variances.

6.1 Introduction

Alzheimer's Disease (AD) is a major type of neurodegenerative brain disease prevalent among the elderly population. The number of individuals affected by AD is approximately 50 million in 2018, with an anticipated threefold increase by 2050 [158]. Timely detection of AD plays a vital role in ensuring appropriate patient care in the context of approval of emerging therapeutics acting on disease course. Biomarkers, including Fluorodeoxyglucose Positron Emission Tomography (FDG-PET) images, have demonstrated effectiveness in early AD diagnosis [6]. Recent years, a lot of studies have focused on utilizing FDG-PET to differentiate AD from Cognitive Control (CN) group, or predict the conversion from Mild Cognitive Impairment (MCI) [206]. While distinguishing between AD and CN groups in FDG-PET brain images is relatively straightforward, differentiating stable MCI (sMCI) from progressive MCI (pMCI) presents a more challenging task due to subtle variations. Even though, diagnosing AD at different stages still plays a crucial role in establishing a more reliable diagnosis and facilitating disease interventions.

FDG-PET based brain networks analysis has proven to be an effective tool for understanding brain function and its alterations under different conditions of AD [100, 239]. The early studies on FDG-PET-based network have primarily focused on group-based analyses which may losing some specific information within an individual brain [228]. Meanwhile, recent studies have highlighted the significance of investigating individual-level brain networks as a promising tool offering valuable insights into personalized brain profiles and their implications for understanding AD [228, 102].

In this study, we aim to contribute to this promising direction by proposing a novel method for constructing individual brain networks from static PET images. More specifically, our contribution can be summarized as below:

- We propose a method for constructing individual brain network from static FDG-PET using Kernel Density Estimation (KDE) to represent brain region's values as a probability, and Dynamic Time Warping (DTW) for computing distance between two probabilities.
- We validate the effectiveness of our proposed method in the context of AD, at multiple condition: CN, sMCI, pMCI and AD.

With this study, our primary objective is to validate the effectiveness of KDE, and DTW in constructing individual brain networks. Additionally, we also aim to study FDG-PET based individual brain connectivity and explore its potential applications in the AD diagnosis. The remainder of this paper is organized as follows: Section 2

describes the used dataset and the proposed method. In Section 3, we present the experimental results and discussions. A conclusion is then given in Section 4.

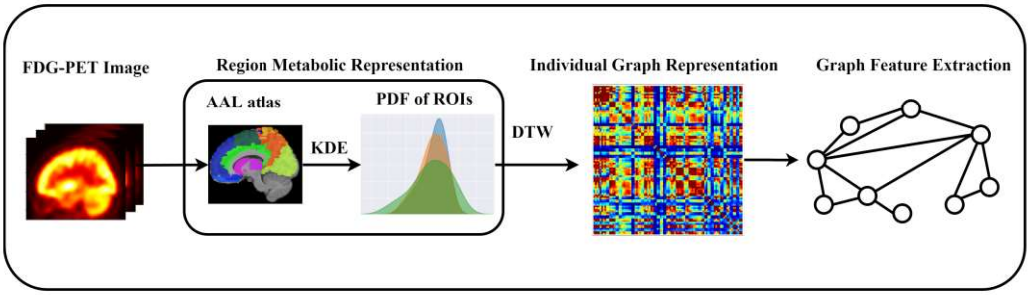

Figure 6.1: Workflow for constructing an individual graph from a static FDG-PET image.

6.2 Materials and Methods

The flowchart of the proposed method, shown in Fig. 6.1, in the following subsections, we introduce the proposed method in details.

Table 6.1: DEMOGRAPHIC AND CLINICAL INFORMATION OF SUBJECT

Characteristic	**CN**	**sMCI**	**pMCI**	**AD**
Number of subjects	242	360	209	237
Female/Male	122/120	153/207	87/122	97/140
Age (Mean ± std.)	73.66 ± 5.66	71.73 ± 7.66	73.89 ± 6.88	75.00 ± 7.91
MMSE (Mean ± std.)	29.03 ± 1.20	28.20 ± 1.59	27.13 ± 1.71	23.19 ± 2.12

6.2.1 Dataset

Data used in the experiment is 18F-FDG PET images downloaded from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. In the ADNI dataset, participants can take several scans at different time points, the first time of scan refers to a baseline scan. Time points after baseline are considered as follow-up time. Data is selected based on following criteria:

- CN - subjects diagnosed as cognitive control at the baseline and do not change over time.
- sMCI - subjects diagnosed as MCI at the baseline and did not convert to AD in the follow-up period at least 24 months.

- pMCI - subjects diagnosed as MCI at the baseline and progressed to AD in the available scan time.
- AD - subjects diagnosed as Alzheimer's disease at the baseline and do not change within the follow-up time.

After downloading the data, images are spatially normalized to MNI template using SPM12 with $2 \times 2 \times 2$ mm voxel size, and $91 \times 109 \times 91$ matrix dimension. After that, images are further smoothed by a Gaussian kernel with a full width at half maximum of 8 mm. All the procedures above are implemented with SPM12 [162]. The intensity normalization is then performed through dividing each voxel intensity by the average value of the cerebellum region. After the preprocessing phase, we have a total 1048 FDG-PET images for later experiments, including 242 CN subjects, 360 sMCI subjects, 209 pMCI subjects, and 237 AD subjects. The demographic and clinical information of subjects is provided in Table 6.1, in which MMSE stands for the Mini-Mental State Examination.

6.2.2 Methods

6.2.2.1 Individual brain network

In order to constructing individual brain networks, we used the KDE and DTW algorithms. More specifically, for constructing an individual metabolic network from each FDG-PET image, we used Talairach atlas [179] to segment brain image into 94 regions of interests (ROIs), these ROIs are considered as nodes for the graph. After that, the intensity values of voxels within each ROI were extracted and used to estimate the probability density function (PDF) of this ROI using KDE method. After that, we used DTW to compute distance between any pair of brain regions, the obtained results is in form of correlation matrix. Based on this metric we construct individual graph for each FDG-PET image. The following sections will be more details about these steps.

KDE method is a non-parametric method used to estimate the probability density function of a random variable based on a set of observed data points [232]. The KDE estimation process involves placing a kernel on each data point and then summing up these kernels to obtain the density estimate. In our study, KDE is applied to each brain region's FDG-PET data to estimate the underlying distribution of glucose metabolism levels within that region at a specific value range and a fixed number of points. By estimating the density of glucose metabolism, KDE enables the representation of brain region activity as probability distributions. This representation allows for more nuanced analysis and facilitates the computation of correlations between brain regions based on their metabolic activity patterns. After that, the DTW algorithm is utilized to compute the similarity between the metabolic activity patterns of all possible pairs of brain regions.

DTW algorithm is a widely used distance measure and alignment algorithm for comparing time series data [188]. The DTW algorithm involves finding the optimal alignment between two time series by minimizing the accumulated distance between

corresponding points. In our case, DTW algorithm is used to compare the KDE's values between two brain regions. Given P and Q are two PDFs defined on the same value ranges, with the same length. The algorithm proceeds as follows:

1. **Input**: Given two PDFs, P and Q of length n.
2. **Initialization**: Create a cost matrix D of size $n \times n$ and initialize all entries to infinity. Set $D[0][0] = 0$.
3. **Dynamic Programming**: For each element $D[i][j]$ in the cost matrix, calculate the cost as the distance between the corresponding elements in P and Q plus the minimum cost from the three neighboring elements: $D[i-1][j-1]$, $D[i][j-1]$, and $D[i-1][j]$.
4. **Optimal Path**: Trace back the minimum cost path from $D[n][n]$ to $D[0][0]$ to obtain the optimal alignment.
5. **Output**: The optimal alignment path represents the matching pairs of points between the two PDFs, and the accumulated cost at $D[n][n]$ represents the similarity between them.

The cost matrix D in the DTW algorithm is computed using the following equation:

$$D[i][j] = d(a_i, b_j) + \min(D[i-1][j-1], D[i][j-1], D[i-1][j])$$

in which $D[i][j]$ represents the cost at position (i, j) in the cost matrix; a_i and b_j are the elements at positions i and j in the PDFs P and Q, respectively; $d(a_i, b_j)$ is the distance measure between a_i and b_j, for simplicity, in our case, we chose the squared Euclidean distance metric.

Finally, the connection between two ROIs (two nodes) was computed by the following equations:

$$c_{DTW}(P,Q) = e^{-DTW(P,Q)}$$

By comparing the metabolic activity patterns using DTW, we obtain a correlation matrix that quantifies the similarity between each pair of brain regions. This correlation matrix serves as the basis for constructing individualized brain networks.

6.2.2.2 Graph Metrics & Graph Comparison

To evaluate the effectiveness of our proposed method, we analyze connectivity matrices from two different perspectives:

1. *Classification*: We use connectivity matrices as features for training Support Vector Machine (SVM) classifiers on various AD classification tasks. To evaluate the effectiveness, we compare the performance between our proposed method and a baseline method. Specifically, we use Kullback-Leibler Similarity (KLS) [228] with same settings of KDE as the comparison method.

2. *Graph Metric Anlysis*: We computed several graph metrics including characteristic path length (Lp), clustering coefficient (Cp), global efficiency (Eg), local efficiency (Eloc). In which, Cp and Eloc are the metrics of functional segregation; Lp and Eg measures functional integration of the brain network, respectively. The definitions and equations of these metrics can be found in [32]. In order to compare the differences in the individual graphs between two group of data, T-test is used to measure difference of individual graph network metrics between two groups. Here, we set a significance level of $p < 0.05$.

6.3 Results and Discussion

6.3.1 Results

Table 6.2: Classification performance on AD classification tasks.

Task	**KLS-based method**				**Our method**			
	ACC	SEN	SPE	AUC	ACC	SEN	SPE	AUC
CN vs. AD	87.93	87.26	88.59	93.97	**88.94**	**87.90**	**89.96**	**95.79**
sMCI vs. pMCI	**69.74**	**56.91**	77.19	**73.62**	69.31	55.69	**77.21**	72.53
pMCI vs. AD	61.46	63.40	59.28	66.27	**63.48**	**65.94**	**60.68**	**68.15**

Table 6.3: Comparing Graph metrics between different groups

group	**Lp**	**Cp**	**Eg**	**Eloc**
CN	2.1080	0.7057	0.4783	0.8079
sMCI	2.1447	0.7012	0.4701	0.8042
pMCI	2.2845	0.6895	0.4429	0.7955
AD	2.4063	0.6727	0.4210	0.7804

We compared the individual networks corresponding to 4 groups: CN, sMCI, pMCI and AD, based on 4 graph measures Lp, Cp, Eg and Eloc (Fig. 6.4 & Table 6.3).

The analysis of individual graph metrics across different conditions revealed substantial differences between the healthy control group and various stages of AD. Notably, a consistent increasing trend was observed in all the metrics (Table 6.3) as disease severity progressed. This observation suggests that these metrics can serve as potential indicators of disease progression and potentially aid in distinguishing between different groups.

Based on the results from statistical tests (Fig. 6.4), it is suggested that characteristic path length can be considered as a good indicator for distinguishing between different conditions.

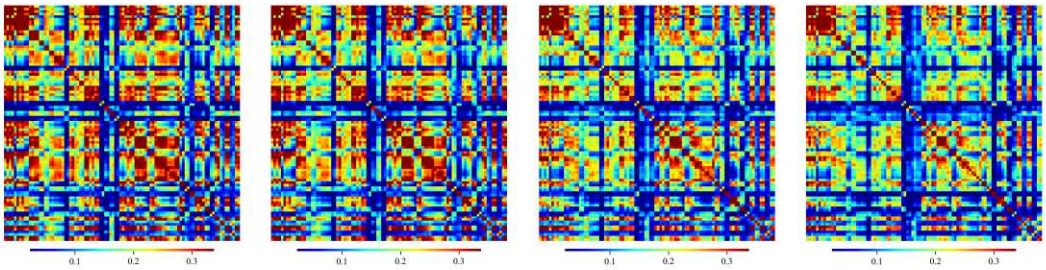

Figure 6.2: Comparing individual graph metrics between different groups

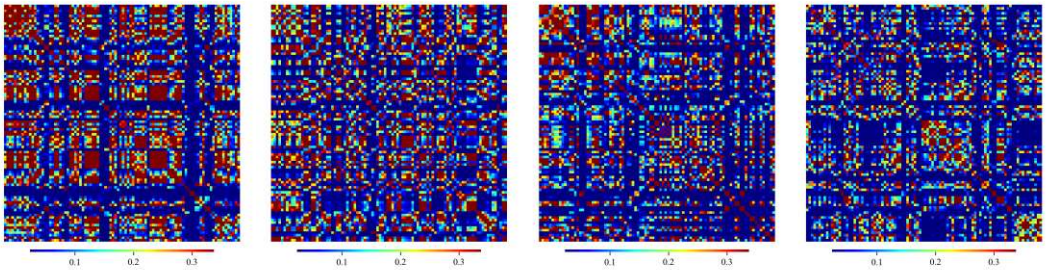

Figure 6.3: Comparing individual graph metrics between different groups

Additionally, metrics Cp and Eloc also exhibited effectiveness in distinguishing between CN, pMCI, and AD (Fig. 6.4). It implies that these metrics can be considered as potential features for characterizing and differentiating disease stages from CN group. In other words, it suggests that for distinguished CN group from AD-related groups we can focus segregation properties of the individual networks.

We also observed that the differences between sMCI and pMCI, as well as pMCI and sMCI, are subtle. This observation is reasonable considering the high similarity between them.

Compared to CN group, from sMCI to pMCI and AD, statistical tests revealed significant differences in all metrics across the groups, with p-values decreasing as disease severity increased (Fig. 6.4). These findings indicate that the observed differences are not random but rather align with the expected progression of the disease and the average MMSE score between groups.

The obtained results demonstrate the effectiveness of the proposed methods in capturing and quantifying the differences in individual brain metabolism network across various conditions of AD. It highlights the potential of our proposed method for clinical applications and disease monitoring.

6.3.2 Discussion

Our results align with prior studies: In our study, both AD and MCI groups exhibited higher characteristic path length values compared to the CN group. This indicates a slower information transmission between nodes in the brain network of disease groups. This finding is consistent with previous studies [128]. The cluster coefficient

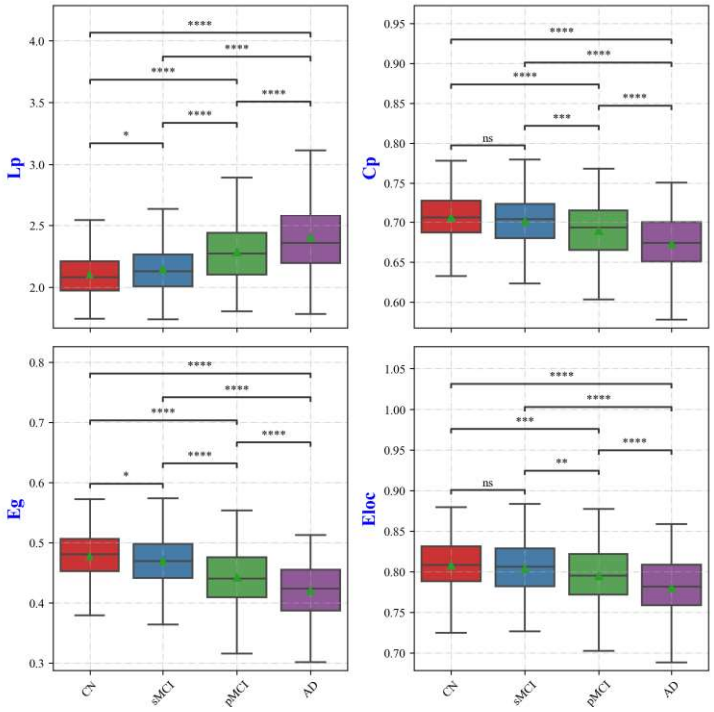

Figure 6.4: Comparing individual graph metrics between different groups

values in the pMCI and AD groups showed a significant increase compared to the CN group, suggesting a tendency towards a more regular network structure. This observation also aligns with the findings reported in [128].

These observations highlight the altered network characteristics in AD and MCI, indicating disruptions in information transmission and the emergence of more regular network patterns. The study findings are in alignment with prior investigations concerning neuron imaging and brain network analysis for AD diagnosis. These results substantiate the efficacy of utilizing KDE and DTW techniques in constructing individual brain networks from FDG-PET images. Our findings provide empirical evidence supporting the concept that individual networks, which consider the intrinsic connectivity patterns within individuals, offer valuable insights into the neurobiological underpinnings of AD. It also highlights the importance of personalized analysis for enhancing diagnostic accuracy in mental health disorders.

Future works: This study demonstrates the efficacy of KDE and DTW for constructing individual brain networks from FDG-PET images. However, to validate the effectiveness of the proposed method, further experiments on different brain diseases and datasets are warranted. Additionally, since the performance of DTW is heavily depending on KDE, it is worthwhile to investigate KDE values at different scales.

Moreover, once the effectiveness of KDE and DTW has been established, leveraging the advantages of DTW becomes feasible. The "warp" property of DTW, allowing for stretching or compressing one signal to match another, enables the comparison of KDE values within the same region of interest (ROI) across different individuals. This

approach is anticipated to elucidate the inter-individual variability within the same brain ROI.

Furthermore, as the current study utilized the Talairach atlas comprising 12 lobes, it is essential to consider brain atlases with a larger number of regions. This way is expected to enhance the granularity and accuracy of the analysis.

6.4 Conclusions

In this study, we have introduced a novel method that incorporates KDE and DTW algorithms for the construction of individual brain networks using FDG-PET images. The effectiveness of our proposed method was validated using the ADNI dataset. By comparing individual graph metrics across the CN, sMCI, pMCI and AD groups, our findings support the presence of disruptions in information transmission within the FDG-PET brain network of AD patients. This study provides evidence for the efficacy of KDE and DTW techniques in constructing individual brain networks from FDG-PET images. Furthermore, our method holds potential for application in other brain disorders and with static FDG-PET images. Future investigations should include additional experiments to further validate the robustness of our approach. Additionally, the consideration of brain atlases with a larger number of regions, as well as exploring KDE at different scales, will facilitate a comprehensive understanding of the limitations and strengths of the proposed method.

6.5 Chapter Conclusion

This chapter collectively highlights and underscores the advantages of individual metabolic networks in enhancing the diagnosis of AD. The methods proposed herein are not limited solely to this dataset but can also be extended to diverse brain disorders. The versatility of these methods opens up broader applications in comprehending and diagnosing various neurological conditions.

Chapter 7

Graph Neural Network Analysis

Summary

7.1 Introduction . . . 138
7.2 Background . . . 139
7.2.1 Related works . . . 139
7.2.2 Graph Notations and Graph Building . . . 140
7.2.3 Graph Neural Networks for Graph Classification . . . 140
7.2.3.1 GATs . . . 141
7.2.3.2 GINs . . . 142
7.2.4 Graph Readout . . . 142
7.3 Data & Experimental Setup . . . 142
7.3.1 Datasets . . . 142
7.3.2 Graph Building . . . 144
7.3.3 Model selection & assessment . . . 145
7.3.4 Non-graph baselines . . . 145
7.4 Results . . . 145
7.4.1 Disorder diagnosis task (dataset) . . . 145
7.4.2 Network Scale . . . 148
7.4.3 Network building method . . . 148
7.4.4 Graph Neural Network Models . . . 151
7.4.5 Interplay between network scales, network building methods and graph neural network models . . . 151
7.5 Discussion . . . 152
7.6 Conclusion . . . 153

In the previous chapter, our focus was on graph analysis using graph metrics. While it revealed many interesting results, there has been a growing concern regarding the use of graph neural networks for studying brain networks. Comparing traditional machine learning with graph metrics to employing graph neural networks (GNNs) for classification reveals several advantages.

Firstly, GNNs enable end-to-end learning, allowing them to directly learn representations from graph-structured data without the need for manual feature engineering. Secondly, GNNs can capture both local and global information encoded

in the graph structure, enabling effective leverage of relationships between nodes and their neighborhoods. Traditional machine learning methods may struggle to capture such complex relationships, particularly in large and intricate graphs. Thirdly, GNNs are inherently adaptable to varying graph structures, automatically adjusting their parameters based on the topology of the input graph. Lastly, GNNs have the capacity to model nonlinear relationships between nodes in the graph, enabling them to capture more intricate patterns and dependencies compared to traditional linear models commonly used with graph metrics.

Despite these advantages, the application of GNNs on static FDG-PET remains limited. There is a need to explore the contribution of graph neural networks for diagnosing brain disorders using FDG-PET images. In this chapter, we extensively compare several common graph neural networks under various network settings for both AD and Depression task classification. This chapter partially addresses the question regarding the contribution of GNNs to diagnosing brain disorders using metabolic networks.

7.1 Introduction

With the advancement of neuroimaging technologies, there has been a growing interest in studying the structural and functional aspects of the brain to elucidate various phenomena related to brain behavior under different conditions, including both healthy and abnormal states[33, 200]. While modalities like structural MRI (sMRI) and functional MRI (fMRI) have been extensively utilized in this regard, Fluorodeoxyglucose-Positron Emission Tomography (FDG-PET) images present a unique perspective as a typical form of nuclear imaging. These images offer valuable information about sugar absorption in the brain, indirectly providing insights into its activity[81]. However, the analysis of FDG-PET data is particularly challenging due to its intricate and complex nature, often characterized by high-dimensional representations. Consequently, there is a pressing need for the development of more effective Artificial Intelligent (AI) analytical tools to fully leverage the potential of FDG-PET imaging in understanding brain function and pathology [71].

Graph-based analysis, particularly graph neural networks (GNNs), have emerged as powerful AI tools for studying Alzheimer's Disease (AD) and other neurological conditions, extensively utilized in fMRI and sMRI diagnosis [243]. However, the application of GNNs to FDG-PET images is relatively limited. In literature, there is very few studies exploring GNNs in fusion FDG-PET images with other modalties. Moreover, when employing metabolic networks and GCN models, several key questions arise, including the optimal number of nodes, defining network connections, network architecture, and training methodologies. As research on the application of FDG-PET images and GCNs for diagnosis remains limited, addressing these questions is crucial for further advancements in FDG-PET image analysis.

To tackle these concerns, this study aims to comprehensively compare the performance of popular GNN models concerning network construction methods and

network scales across different datasets. Specifically, we focus on the ADNI dataset for AD diagnosis, where distinctions between healthy and patient subjects are more evident, and a local dataset featuring patients with depression disorders, where differentiating between healthy and patient subjects poses a greater challenge. To ensure a fair comparison, baseline experiments are also conducted using non-graph methods. In summary, in this work, we:

1. Compare the performance of different GNN models, including Graph Convolutional Network (GCN), Graph Attention Network (GAT), and Graph Isomorphism Network (GIN) models on FDG-PET images.

2. Validate the effectiveness of each GNNs model on different datasets (Alzheimer's Disease Neuroimaging Initiative (ADNI), local dataset) with various settings, including methods for building graph (Dynamic Time Warping (DTW), Kullback-Leibler Similarity (KLS)), scales (Schaefer atlases with 100, 200, 300 regions).

3. Compare graph-based methods with non-graph-based methods, including Support Vector Machine (SVM) and Multi-Layer Perceptron (MLP).

Through this, we aim to provide insights into empirical analysis of different GNNs in applications on FDG-PET brain networks.

The remainder of this paper is organized as follows: in Section 7.2, related work is briefly described, and three types of GNNs are introduced. In Section 7.3, we provide an overview of the dataset and experimental setups. The experimental results are presented in Section 7.4. Finally, a discussion and conclusions are drawn in Sections 7.5 and 7.6, respectively.

7.2 Background

7.2.1 Related works

FDG-PET imaging has become a standard tool in diagnosing various brain disorders, including Alzheimer's disease, Parkinson's disease, and others [81]. Its ability to provide insights into cerebral glucose metabolism offers valuable information for understanding neurological conditions. However, the complex nature of FDG-PET data necessitates effective AI analytical tools for accurate diagnosis and interpretation [71]

Graph-based approaches have emerged as promising tools for analyzing FDG-PET data. While many studies have focused on traditional graph metrics analysis, the application of GNNs in neuroimaging, particularly FDG-PET, remains relatively underexplored. Previous research has primarily utilized GNNs in other neuroimaging modalities such fMRI, sMRI [243]. The limited number of studies on the application of GNNs to FDG-PET raises questions about their potential contributions to the diagnosis of brain disorders using this imaging modality.

Inspired by benchmarking studies of GNNs in other domains such as fMRI [65] and gene data analysis, our study aims to address this gap by conducting a comprehensive benchmarking analysis. We focus on three widely used GNN architectures: GCN, GIN, GAT. Additionally, we include comparisons with non-graph methods to assess the effectiveness of GNNs in FDG-PET analysis. Beyond evaluating different network architectures, we also explore various settings such as network scale and network building methods. This comprehensive approach allows us to gain insights into the performance of GNNs under different conditions and datasets.

Through this work, we aim to shed light on the potential contributions of GNNs to FDG-PET-based diagnosis of brain disorders. Additionally, we provide recommendations for researchers regarding the considerations and challenges associated with employing GNNs in FDG-PET analysis.

7.2.2 Graph Notations and Graph Building

As this work focus on static FDG-PET image, we only focus on static representation of graph, in other word we assume no temporal evolution of the graph structure and node features. A conventional graph is normally represented as a combination of $G = (V, E)$ in which V is set of nodes, E is set of edges. For definining the nodes, the most common approache is to using predefined brain atlas such as Schaefer, Automated Anatomical Labeling to seperating brain images into distinguished regions and treating those regions as nodes for graph. Then, the edges is defining as correlation between nodes. After constructing graph, then graph metrics can be computed for study topology of graph. Even graph metrics is common approach, they are limited in the choice of graph metrics, as well as manully defining. Meanwhile, recently, there is a growing intrest in using graph neural network for autmatically extracting graph feature and learning represntation from data, as can be found in [**1**, **2**, **3**]. With the development of neural network and graph neural network, different graph architecture have been proposed and development. Inspired from previous work with fMRI, in this work, we focus on three type of graph neural network: graph convolutional neural network, GAT, GIN models. These types of network are selected based on ... A short briefly description of those models is provided in section 7.2.3, for fully description can be found in previous survey [**4**, **5**, **6**].

7.2.3 Graph Neural Networks for Graph Classification

For the task of graph classification, the objective of a GNN is to learn a function denoted as $f(\cdot)$, which maps a given graph to a feature vector. This feature vector encapsulates essential information about the graph necessary for classification. Mathematically, the forward function can be represented as:

$$y = g(f(G))$$

Here, $y \in \mathbb{R}$ represents the label of the graph, and $g(\cdot)$ signifies a function parameter-

ized by a neural network that maps the output feature vector of the GNN to the graph's label.

Graph Convolutional Networks (GCNs)

Graph Convolutional Networks (GCNs) learn representations of graphs by approximating spectral convolutions. In spectral convolutions, the graph signal x is multiplied by a filter s_θ in the Fourier domain, given by:

$$s_\theta * x = U s_\theta U^T x$$

Here, U represents the matrix of eigenvectors of the normalized graph Laplacian L, where $L = I - D^{-1/2} A D^{-1/2} = U \Lambda U^T$, and Λ is a diagonal matrix of eigenvalues. To simplify the computational complexity, GCNs utilize a truncated expansion of Chebyshev polynomials up to a certain order K. This approximation allows GCNs to effectively convolve the graph signal with the filter.

The layer-wise propagation rule of GCNs in the spatial domain is defined as:

$$H^{k+1} = \sigma(\hat{D}^{-1/2} \hat{A} \hat{D}^{-1/2} H^k W^k)$$

Here, H^k denotes the graph signal representation at layer k, with $H(0) = X$. $\hat{A}$ is the adjacency matrix with self-connections, $\hat{D}$ represents the degree matrix, W^k is a trainable weight matrix, and σ denotes a non-linear activation function. This equation represents the message passing and update operations on the nodes in a GCN. More details can be found in [**1**, **2**, **3**, **4**]

7.2.3.1 GATs

The Graph Attention Network (GAT) differs from a Graph Convolutional Network (GCN) in its update function, which combines node features h_i with aggregated neighborhood features $h_{\mathcal{N}_i}$ using a permutation-invariant mechanism. Unlike GCNs, where the importance of neighboring nodes is defined by c_{ij}, the normalized adjacency matrix entry $\hat{A}_{ij}$, GATs utilize a self-attention mechanism to implicitly learn node importance.

Formally, the update rule for each node i in a GAT layer can be expressed as:

$$h_i^{k+1} = \phi(h_i^k, \sum_{j \in \mathcal{N}_i} a(h_i^k, h_j^k) \psi(h_j^k))$$

$$a(h_i^k, h_j^k) = Softmax(\sigma(m^T [W h_i^k || W h_j^k]))$$

Here, $||$ denotes concatenation, σ is a non-linear activation function, and m represents a trainable single-layer feed-forward neural network. The attention mechanism can be extended to a multi-head attention by learning M different copies of $a(h_i^k, h_j^k)$ and computing their average.

7.2.3.2 GINs

The Graph Isomorphism Network (GIN) is a variant of GNN specifically designed for graph classification tasks. Unlike GCNs and GATs, which employ various permutation-invariant functions to aggregate neighborhood features, GINs typically utilize only the summation operation. Additionally, GINs do not incorporate a transformation function ψ before combining neighborhood features with original node features. Instead, a multi-layer perceptron (MLP) is directly applied to the concatenated features.

The update rule for each node in a GIN layer is expressed as:

$$h_i^{k+1} = MLP\left((1+\epsilon)h_i^k + \sum_{j \in \mathcal{N}_i} h_j^k\right)$$

Here, ϵ represents a learnable parameter initialized to zero, determining the amount of original information retained in the node for the subsequent representation.

7.2.4 Graph Readout

Following feature extraction using any of the proposed GNN architectures, the readout phase computes a feature vector for the whole graph using some permutation-invariant readout function to be passed as an input to the final classification layer. In this work we experiment with 3 different permutation-invariant functions to summarize the graph features after the last graph layer. Namely *Mean()*, *Mean()||Max()* and *Sum()*. The selection of which function to use for each architecture/dataset combination is set as a hyperparameter. We empirically observe that *Mean()* and *Mean()||Max()* provide the most stable results to all GNN architectures with exception of the GIN architecture where the *Sum()* operation marginally improve the performance of the model.

7.3 Data & Experimental Setup

Fig 8.1 despict flow chart of the experiment. In next subsection, we will briefly about dataset and substeps for building and training graph neural networks.

7.3.1 Datasets

Throughout this study we evaluate the performance of the models on two different datasets. Specifically, experiments are performed on ADNI datasets and local dataset. While with ADNI dataset, the differences between Alzheimer's disease and healthy control groups is well-observation, the local dataset regarding to psychiatric with subtle changes between depression patients and control groups. Belows are short description about dataset, including data selection and preprocessing steps:

ADNI dataset Data used in the experiment is 18F-FDG PET images downloaded from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. In the ADNI

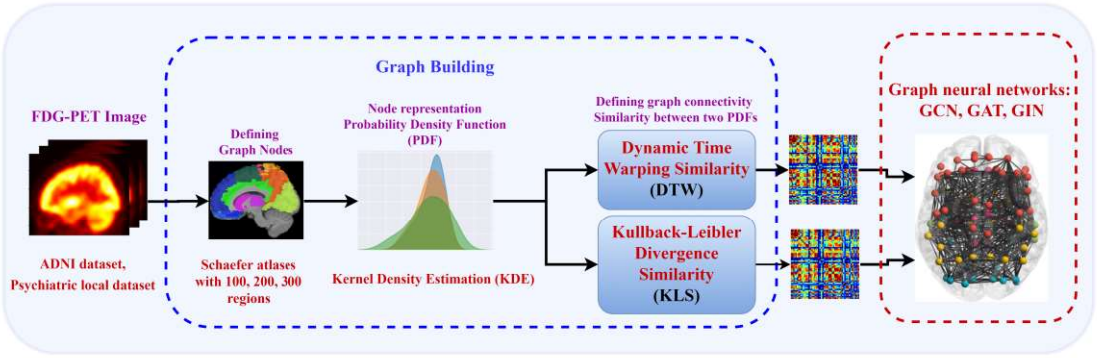

Figure 7.1: Pipeline of experiments.

dataset, participants can take several scans at different time points, the first time of scan refers to a baseline scan. Time points after baseline are considered as follow-up time. Data is selected based on following criteria:

- CN - subjects diagnosed as cognitive control at the baseline and do not change over time.
- AD - subjects diagnosed as Alzheimer's disease at the baseline and do not change within the follow-up time.

After downloading the data, images are spatially normalized to MNI template using SPM12 with $2 \times 2 \times 2$ mm voxel size, and $91 \times 109 \times 91$ matrix dimension. After that, images are further smoothed by a Gaussian kernel with a full width at half maximum of 8 mm. All the procedures above are implemented with SPM12 [161]. The intensity normalization is then performed through dividing each voxel intensity by the average value of the cerebellum region. After the preprocessing phase, we have a total 1048 FDG-PET images for later experiments, including 242 CN subjects, 360 sMCI subjects, 209 pMCI subjects, and 237 AD subjects. The demographic and clinical information of subjects is provided in Table 7.1, in which MMSE stands for the Mini-Mental State Examination.

Local dataset The local dataset utilized in this study consists of FDG-PET images collected from patients at Hospital A, following rigorous adherence to standard acquisition protocols. These images were obtained as part of routine clinical assessments, providing a valuable resource for investigating psychiatric conditions. The preprocessing pipeline for these FDG-PET images involved several critical steps aimed at enhancing data quality and ensuring comparability across subjects. These steps included motion correction to mitigate artifacts caused by patient movement, spatial normalization to align images to a common anatomical template for consistent spatial interpretation, intensity normalization to account for variations in scanner settings, and potentially the application of noise reduction techniques to improve signal-to-noise ratio. By implementing these preprocessing procedures, the local dataset was prepared for subsequent analysis, facilitating robust and reliable investigations into

the neurobiological underpinnings of psychiatric disorders. The demographic for ADNI, and psychiatric dataset can be found in below tables:

Table 7.1: DEMOGRAPHIC AND CLINICAL INFORMATION OF SUBJECT

Characteristic	CN	AD
Number of subjects	242	237
Female/Male	122/120	97/140
Age (Mean ± std.)	73.66 ± 5.66	75.00 ± 7.91
MMSE (Mean ± std.)	29.03 ± 1.20	23.19 ± 2.12

7.3.2 Graph Building

After preprocessing steps, we use use Schaefer atlases [186], in order to seperate brain into distinguished brain regions, and treating those regions as nodes. In order to examinate the affect of network scales to results we experiments with Schafer atlas at 3 scales: 100, 200 and 300 regions. For defining connections (edges), we follow two approaches: KLS-based and DTW as proposed in [229], [213] respectively. Note that, while KLS based on information theory, DTW is based on distance comparisons. After that, each graph is represented as an adjency matrix. Below we short briefly describe about KLS and DTW methods.

KDE method is a non-parametric method used to estimate the probability density function of a random variable based on a set of observed data points [233]. The KDE estimation process involves placing a kernel on each data point and then summing up these kernels to obtain the density estimate. In our study, KDE is applied to each brain region's FDG-PET data to estimate the underlying distribution of glucose metabolism levels within that region at a specific value range and a fixed number of points. By estimating the density of glucose metabolism, KDE enables the representation of brain region activity as probability distributions. This representation allows for more nuanced analysis and facilitates the computation of correlations between brain regions based on their metabolic activity patterns. After that, the similarity between two nodes (connections) is defined based on computing the similarity between the metabolic activity patterns of all possible pairs of brain regions.

In order to compare the PDFs between a pair of brain regions, in this work, we explore two methods for comparing PDFs: Kullback-Leibler Divergence (KLD) and Dynamic Time Warping (DTW).

Kullback-Leibler Divergence (KLD) is a fundamental concept in information theory, measuring the relative entropy between two probability distributions. For discrete PDFs P and Q, KLD is computed as:

$$KLD(P||Q) = \sum_i P(i) \log \frac{P(i)}{Q(i)}$$

KLD provides a direct assessment of the difference in information content between two distributions. However, it assumes independence between data points and may not capture local similarities effectively.

Dynamic Time Warping (DTW) is a versatile technique commonly used to compare sequences with varying temporal characteristics. When applied to PDFs, DTW aligns the distributions in a nonlinear time warping manner, facilitating optimal matching between elements. The DTW distance between PDFs P and Q is given by:

$$DTW(P,Q) = \min\left(\sqrt{\sum_{i=1}^{N}(P(i)-Q(j))^2}\right)$$

7.3.3 Model selection & assessment

In order to conduct experiments, we use the optimal parameters from previous work [65] and split the data into train/validation/test sets with a ratio of 3:1:1. This process is repeated 5 times, and the average results are reported. (See Supplementary Information for details about training the GNN models and their hyperparameters.)

7.3.4 Non-graph baselines

To ensure a fair assessment of whether GNNs can effectively leverage graph structures, it is necessary to evaluate their performance against baselines that lack inductive awareness of the data's structure. A comparable framework should also be applied to hyperparameter search and model selection. In this study, two distinct structure-agnostic baselines are employed. Specifically, we utilize a multi-layer perceptron (MLP) and support vector machines with a radial basis function kernel (SVM-rbf) as baselines for learning from static representations. (For further information on training the baselines and their hyperparameters, please refer to the Supplementary Information.)

7.4 Results

7.4.1 Disorder diagnosis task (dataset)

Fig. 7.2 and Table 7.2, 7.3 present a comparative analysis of the average performance of baseline methods, including MLP and RBF SVM, alongside graph neural networks such as GCN, GIN, and GAT, across two distinct datasets: ADNI and the psychiatric dataset.

Across various experimental settings, it is evident that the performance on the ADNI dataset consistently surpasses that of the psychiatric dataset across all metrics, including ACC, SEN, SPE for almost type of classification models. This discrepancy in performance can be attributed to the inherently clearer distinction between AD and CN individuals within the ADNI dataset, particularly evident in the analysis of

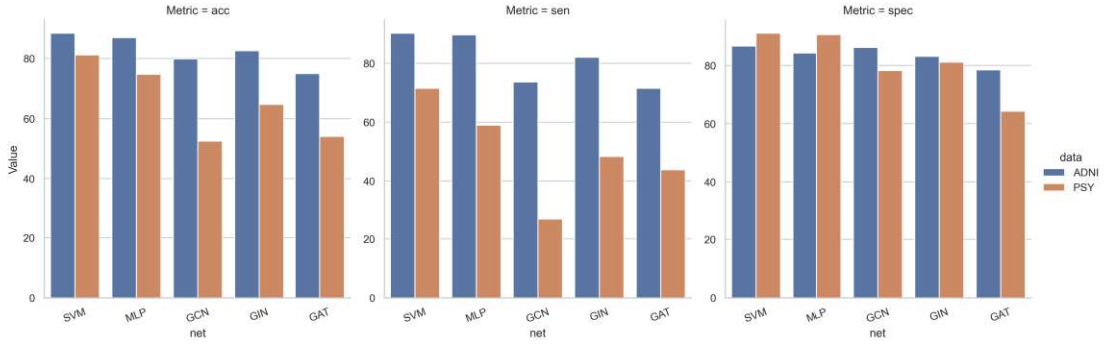

Figure 7.2: Comparison of average performance of baseline methods (MLP, RBF SVM) with GNNs (GCN, GIN, GAT) on ADNI and psychiatric datasets across various metrics.

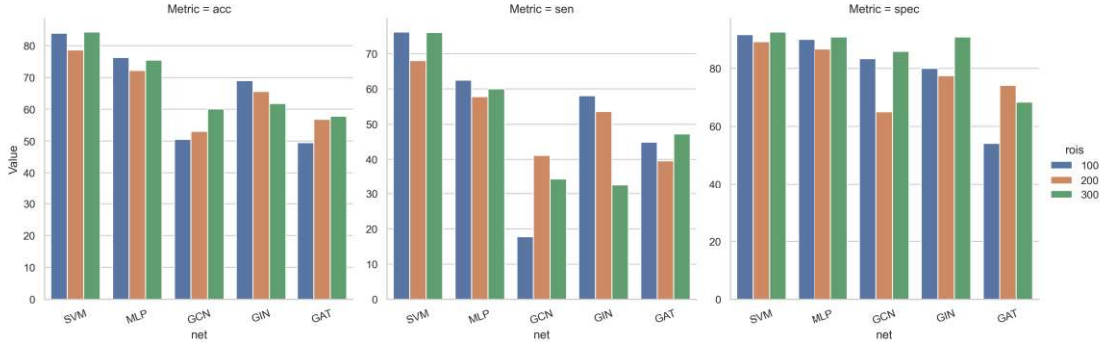

Figure 7.3: Performance comparison of all methods across different scales (100, 200, 300 regions) on the psychiatric dataset.

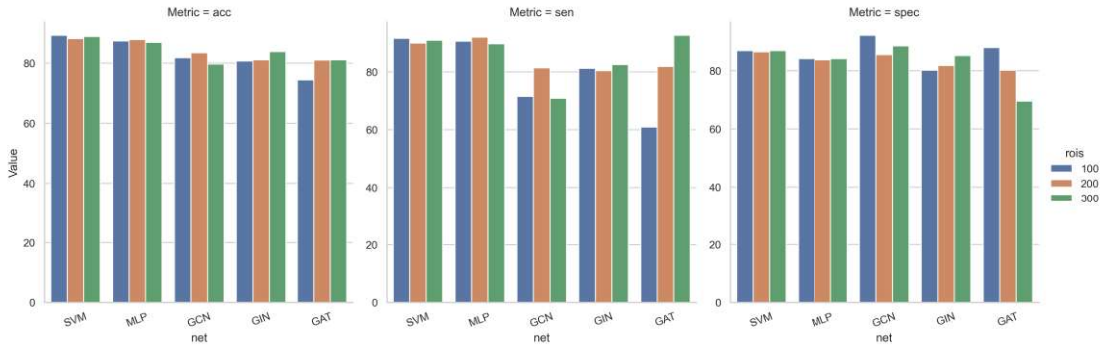

Figure 7.4: Performance comparison of all methods across different scales (100, 200, 300 regions) on the ADNI dataset.

FDG-PET images. Conversely, the psychiatric dataset, focusing on depression and CN individuals, presents a more challenging classification task, resulting in comparatively lower performance across all models. This observation suggests that the characteristic of data significantly contributed to the diagnosis performance of all models.

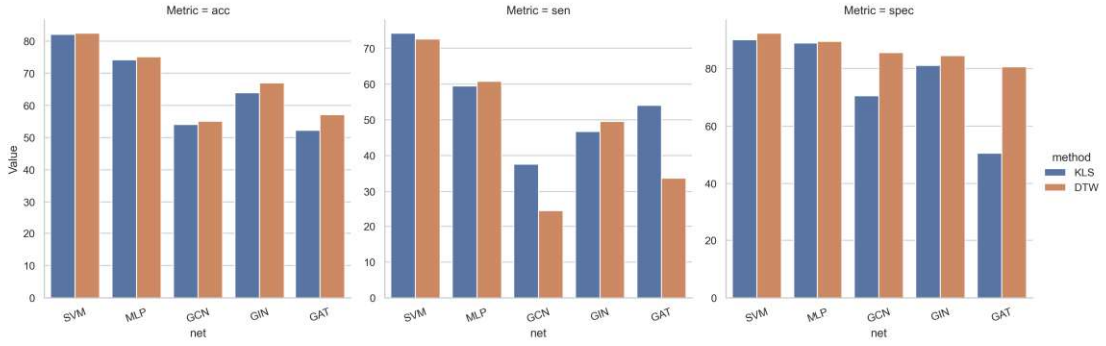

Figure 7.5: Performance comparison of all methods when comparing two graph building methods (DTW, KLS) on the psychiatric dataset.

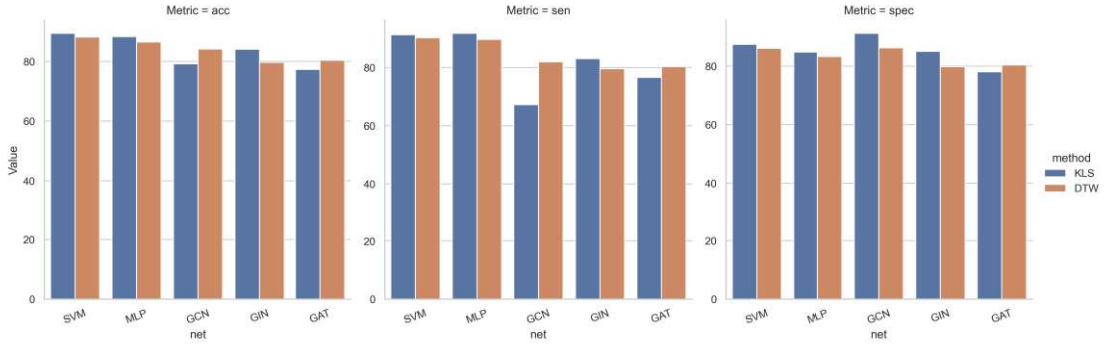

Figure 7.6: Performance comparison of all methods when comparing two graph building methods (DTW, KLS) on the ADNI dataset.

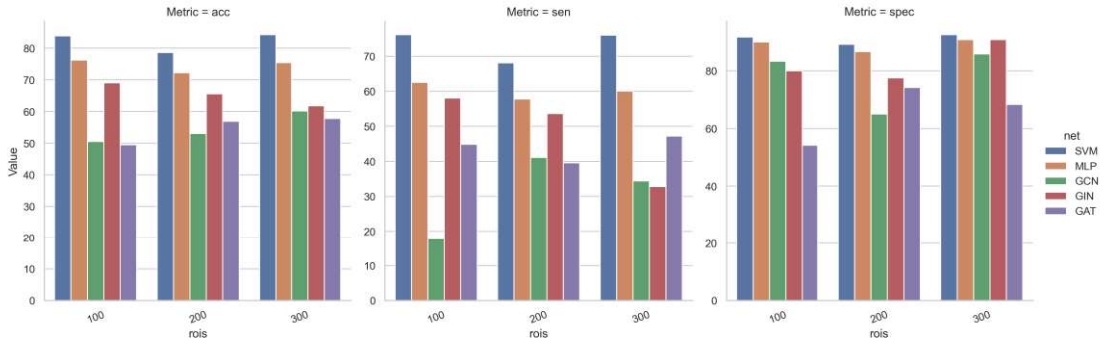

Figure 7.7: Performance comparison of all methods when graph neural network structure (GCN, GIN, GAT) on the psychiatric dataset.

In general, based on Fig. 7.2 we also observe that SVM and MLP models demonstrate superior performance across both datasets. Among the GNNs, the GIN architecture tends to outperform GCN and GATn Notably, GCN exhibits the lowest performance on the psychiatric dataset, particularly in terms of ACC and SEN. The various performance between GNN models suggests that each model has it own capability to learn useful

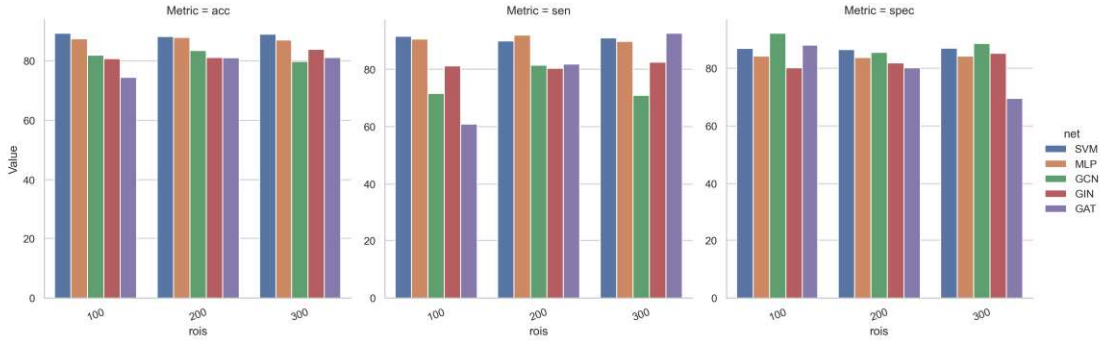

Figure 7.8: Performance comparison of all methods when graph neural network structure (GCN, GIN, GAT) on the ADNI dataset.

representation from data, and it is important to make some experiments to select the most suitable one. In the next, subsection, we will go to more detail analysis effect of network scale, network building methods and graph neural network architectures on each dataset separately.

7.4.2 Network Scale

In Fig. 7.3 and 7.4, we analyze the performance of all methods across varying scales (100, 200, 300 regions) corresponding to the psychiatric and ADNI datasets, respectively.

On the psychiatric dataset, it is notable that as the scales change, the performance of the methods exhibits variability. We can observe that for GCN, ACC and SEN is improved as increasing scale, this observation is also true for ACC of GAT mode. Eventhough, a discernible trend is not readily apparent. For baseline models, despite fluctuations in performance, there is no clear indication of improvement or deterioration as scale increases. This suggests a certain level of inconsistency in the relationship between scale and model performance within the context of the psychiatric dataset.

Conversely, on the ADNI dataset, there is a discernible pattern wherein increasing scales correspond to improved performance, particularly for the GNN models, with GIN showing notable enhancement for almost metrics. In contrast, the baseline methods, including MLP and SVM, do not exhibit significant performance improvements with increasing scales. This observation suggests that GNN models may possess a unique capability to capitalize on larger networks, thereby enhancing their performance in tasks characterized by more extensive and complex data representations, such as the ADNI dataset.

7.4.3 Network building method

Fig. 7.5 and 7.6 present a comparative analysis of the performance of all classifier models when employing two distinct graph building methods, Dynamic Time Warping

Table 7.2: Comparison of Performance Metrics for various models on ADNI and Psychiatric Local Dataset with a network proportion of 0.2.

Models			***ADNI***			***Psychiatric Local Dataset***		
			ACC (%)	SEN (%)	SPE (%)	ACC (%)	SEN (%)	SPE (%)
DTW	100	GCN	73.76	71.06	76.48	48.04	17.78	78.34
		GIN	77.66	77.18	78.10	54.04	41.36	66.66
		GAT	65.56	66.56	64.62	50.00	20.00	80.00
		SVM	84.72	87.14	82.28	83.76	72.52	95.00
		MLP	83.26	85.10	81.44	73.04	54.44	91.68
	200	GCN	76.96	70.80	83.14	50.42	2.50	98.34
		GIN	81.20	82.24	80.14	75.84	68.34	83.34
		GAT	60.06	54.36	65.74	54.72	62.78	46.66
		SVM	86.62	88.44	84.80	75.14	58.62	91.68
		MLP	84.52	88.86	80.16	73.90	54.44	93.34
	300	GCN	76.70	75.86	77.54	50.56	24.44	76.68
		GIN	84.60	84.76	84.42	69.46	57.22	81.66
		GAT	69.00	69.68	68.30	57.78	55.56	60.00
		SVM	88.70	88.86	88.62	81.66	68.34	95.02
		MLP	86.00	86.36	85.70	78.08	59.46	96.68
KLS	100	GCN	77.66	57.00	98.38	48.60	35.56	61.66
		GIN	84.96	84.80	85.12	56.54	23.04	90.00
		GAT	79.60	64.12	95.08	54.02	23.04	85.00
		SVM	90.38	90.94	89.84	79.86	73.08	86.68
		MLP	89.98	91.36	88.62	79.72	66.12	93.34
	200	GCN	85.34	80.56	90.10	49.44	2.22	96.66
		GIN	85.94	83.08	88.84	69.04	56.40	81.66
		GAT	80.14	70.58	89.68	53.74	60.84	46.68
		SVM	89.52	92.96	86.12	79.72	72.80	86.66
		MLP	88.72	91.74	85.68	74.58	60.86	88.32
	300	GCN	78.16	80.74	75.60	55.14	51.94	58.34
		GIN	85.16	83.94	86.38	59.02	44.74	73.34
		GAT	72.00	61.76	82.24	50.00	40.00	60.00
		SVM	88.30	89.28	87.34	81.10	72.24	90.00
		MLP	86.44	88.08	84.84	70.00	51.70	88.36

and Kullback-Leibler Similarity, corresponding to the psychiatric and ADNI datasets, respectively.

On the psychiatric dataset, it is clearly that DTW consistently yields superior performance in terms of ACC and SEN across almost classifier models compared to KLS. This observation underscores the efficacy of DTW, particularly its ability to control window size, allowing it to capture subtle changes more effectively. However, on the ADNI dataset, while DTW still demonstrates favorable performance, the difference between

Table 7.3: Comparison of Performance Metrics for various models on ADNI and Psychiatric Local Dataset with a network proportion of 0.5.

Models			***ADNI***			***Psychiatric Local Dataset***		
			ACC (%)	SEN (%)	SPE (%)	ACC (%)	SEN (%)	SPE (%)
DTW	100	GCN	83.04	78.48	87.62	51.38	11.12	91.66
		GIN	77.24	78.02	76.46	69.72	59.46	79.98
		GAT	73.90	63.10	84.72	49.30	18.60	80.00
		SVM	88.52	91.30	85.68	85.56	79.46	91.68
		MLP	85.80	87.62	83.98	78.46	68.64	88.34
	200	GCN	81.80	78.04	85.52	48.88	22.78	75.00
		GIN	78.54	79.36	77.68	63.48	50.28	76.68
		GAT	83.78	82.76	84.80	59.02	41.38	76.68
		SVM	86.82	88.86	84.80	75.54	61.12	90.02
		MLP	86.36	91.30	81.42	71.12	52.22	90.02
	300	GCN	87.48	89.24	85.70	65.00	40.00	90.00
		GIN	83.32	81.44	85.26	67.80	38.88	96.66
		GAT	83.36	95.04	71.68	63.06	41.12	85.00
		SVM	89.12	90.50	87.78	86.12	77.24	95.02
		MLP	87.26	90.10	84.42	75.70	61.42	90.02
KLS	100	GCN	80.62	64.58	96.70	49.72	24.44	75.00
		GIN	84.16	84.38	83.92	68.32	56.68	79.98
		GAT	74.98	58.72	91.26	49.72	71.12	28.34
		SVM	89.96	91.76	88.18	82.22	72.80	91.68
		MLP	88.92	93.42	84.38	74.04	56.40	91.66
	200	GCN	85.10	84.74	85.48	57.24	59.44	55.00
		GIN	83.66	81.36	85.98	67.64	56.94	78.34
		GAT	78.22	80.92	75.52	54.72	37.76	71.66
		SVM	89.52	90.88	88.18	81.68	75.02	88.34
		MLP	89.32	92.56	86.08	73.32	63.36	83.34
	300	GCN	71.96	52.50	91.42	55.28	28.88	81.66
		GIN	84.34	83.50	85.14	55.84	26.68	85.00
		GAT	78.78	90.08	67.44	52.50	53.34	51.68
		SVM	88.68	91.32	86.10	82.36	74.74	90.00
		MLP	86.62	89.28	83.98	75.16	58.64	91.66

DTW and KLS is less pronounced, with KLS showing a slightly improved performance. This nuanced contrast suggests that while DTW maintains its effectiveness, KLS may exhibit certain advantages in specific contexts within the ADNI dataset.

7.4.4 Graph Neural Network Models

From Table 7.3 and Fig. 7.3, 7.4, across both datasets and network scales, we found that the performance of GAN, GCN, and GIN models varied. No single model consistently outperformed the others across all tasks and scales. However, in general, GIN seem to be the better one when gave a better performance on psychiatric dataset (Fig. 7.3), and more stable performance on ADNI dataset (Fig. 7.4). Instead, the choice of model may depend on task-specific requirements and the characteristics of the dataset.

Comparing performance of GNN models to baseline models, in general, the average performance is lower in almost cases. However, the difference it not large, in some cases, the performance of GNNs is very competitive compared to baseslines. Moreover, it is important to note that, in this experiments designs we only training GNNs with small number of epochs with fully connected connection between all layers. There is a large room for finetuning networks to improve the performance. Lastly, as mentioned in previous subsection, GNN models seem to scale well with the network scales. All of this suggested for the high potential of GNNs for better performance and improving diagnosis of brain disorder using FDG-PET images, especially network at large scales.

7.4.5 Interplay between network scales, network building methods and graph neural network models

Our analysis highlights the complex interplay between graph network models, tasks, network scales, and network building methods. Overall, these findings underscore the importance of selecting an appropriate graph building method tailored to the characteristics of the dataset and the task at hand. The superior performance of DTW on the psychiatric dataset suggests its efficacy in capturing nuanced variations and highlights its potential utility in tasks requiring sensitivity to subtle changes.

Specifically, based on the comparisons made across Fig. 7.2 to 7.8, several trends emerge, highlighting the significance of network scale, graph building method, and network architecture in the context of the studied datasets. All of the observations seem to suggest that when designing models for classification tasks on datasets, researchers should consider the following:

- **Network Scale**: Larger networks may be beneficial for tasks requiring complex data representations, such as ADNI classification. However, the impact of network scale may vary across datasets.
- **Graph Building Method**: DTW shows promise in capturing subtle changes within the data, but the choice between DTW and other methods should be guided by the specific characteristics of the dataset.
- **Network Architecture**: Both traditional methods like SVM and MLP, as well as modern GNN architectures, specifically GINs, offer viable options. The selection should be based on the dataset's characteristics and the task requirements, with a preference for GNNs in scenarios involving extensive network connectivity.

7.5 Discussion

In this study, we conducted a comprehensive analysis of the performance of various graph network models across different tasks, network scales, and network building methods. Our findings shed light on the nuanced interplay between model architectures, dataset characteristics, and methods for building network in graph-based classification tasks

Contribution of the Study and Main Findings: This study contributes to the field of FDG-PET disorders analysis by evaluating the performance of various machine learning models, including GNNs, on two distinct datasets: ADNI and a psychiatric dataset. The main findings highlight the potential of GNNs in enhancing classification performance, particularly in tasks involving complex data representations. Specifically, our analysis reveals that while GNN models initially exhibit lower performance compared to baseline methods such as SVM and MLP, they demonstrate scalability and the ability to improve performance with larger network scales. Moreover, among the GNN architectures investigated, the graph isomorphism network emerges as a promising model, showing superior performance and scalability. Importantly, we provide public access to the code used in our experiments, facilitating further exploration and validation of GNNs in FDG-PET disorders analysis.

Contribution to the Research Community: Our study contributes to the research community by providing publicly available code for exploring GNNs in the context of FDG-PET disorders analysis. This resource facilitates reproducibility and encourages collaboration, enabling researchers to build upon our findings and investigate novel approaches for disease classification. Moreover, the framework developed in this study serves as a foundation for future research endeavors, offering a valuable tool for evaluating and comparing different machine learning models in similar applications. By sharing our code and methodology, we aim to foster innovation and advance the understanding of neurodegenerative and psychiatric disorders through machine learning techniques.

Limitations and Future Work: Despite the significant insights gained from this study, several limitations warrant consideration. Firstly, the evaluation is restricted to a limited set of GNN architectures, and further investigation into novel architectures may yield additional insights. Additionally, the study focuses on FDG-PET imaging data, and future work could explore the application of GNNs to other modalities or multimodal datasets for comprehensive analysis. Furthermore, the impact of hyperparameter tuning and optimization techniques on GNN performance warrants exploration to enhance model robustness and generalization. Additionally, the lower performance of GNNs compared to baseline methods does not imply inferiority but underscores the importance of careful experimental design, including the choice of graph construction method and network parameters. It is essential to acknowledge the influence of these factors and optimize them to achieve optimal GNN performance. Moreover, future research should prioritize reproducibility of results and explore methods for enhancing the explainability of GNN models.

7.6 Conclusion

In this study, we conducted a thorough analysis of the performance of various graph network models across different tasks, network scales, and network building methods in the analysis of FDG-PET imaging data for neurodegenerative and psychiatric disorders. Our findings underscore the potential of GNNs in enhancing classification performance, especially when dealing with complex data representations. While GNNs initially exhibit lower performance compared to traditional baseline methods, careful experimental design, including the choice of graph construction method and network parameters, can significantly improve their efficacy. Notably, the graph isomorphism network emerges as a promising GNN architecture, demonstrating superior scalability and performance across different network scales. Our findings offer valuable insights into the complex interplay betwen model architecture, dataset characteristics, and methods for building network in graph-based classification tasks.

For the future work, the need for further exploration of other GNN architectures, the application of GNNs to other modalities or multimodal datasets, and the optimization of hyperparameters and training techniques to enhance model robustness and generalization, or explore methods for enhancing the explainability of GNN models, should be considered as interesting approaches.

Last but not least, this study not only advances our understanding of the potential of GNNs in FDG-PET disorders analysis but also provides a framework for future research endeavors in this field. By sharing our code and methodology, we hope to contribute to the development of more effective models for the diagnosis and prognosis of neurodegenerative and psychiatric disorders.

Chapter 8

Enhancing Depression Diagnosis Using FDG-PET Images with HyperGraphs

Summary

8.1 Introduction . . . 155
8.2 Materials and Methods . . . 156
8.2.1 Datasets . . . 156
8.2.2 Methods . . . 157
8.2.2.1 Graph and Hypergraph Annotations . . . 157
8.2.2.2 Pairwise Graph and Hypergraph Building . . . 157
8.2.2.3 Graph Neural Networks . . . 158
8.2.3 Experimental settings . . . 159
8.2.3.1 Network structure . . . 159
8.2.3.2 Network training . . . 159
8.2.3.3 Performance Evaluation . . . 159
8.3 Results and Discussion . . . 160
8.3.1 Results . . . 160
8.3.2 Discussion . . . 161
8.4 Conclusion . . . 162

The previous chapter demonstrated the advantages of GNNs for analyzing brain disorders. It becomes evident, particularly in psychiatric disorders like Depression, where the distinction between the healthy control and patient groups is not significant, and changes are subtle, the overall performance of GNNs tends to be slightly lower compared to AD disorders. To achieve better performance in such cases, a more powerful framework is required. Hypergraphs offer promising solutions in this regard.

Hypergraphs provide distinct advantages over conventional graphs for capturing complex patterns in brain network disorders. Firstly, hypergraphs can represent higher-order relationships among multiple nodes simultaneously, unlike conventional graphs that only capture pairwise relationships. In brain network disorders

characterized by interactions involving multiple nodes concurrently, hypergraphs provide a more accurate representation of these intricate relationships. Secondly, hypergraphs excel at capturing complex dependencies and interactions among nodes, including higher-order dependencies that may not be adequately captured by conventional graphs. In brain network disorders characterized by intricate patterns of connectivity and dependencies among brain regions, hypergraphs offer a more nuanced representation of these complexities.

Therefore, in this chapter, we conduct a simple comparison of the performance of graph neural networks on hypergraphs and conventional graphs. This chapter emphasizes the contribution of graph neural networks as well as the promising potential of hypergraphs for accurately characterizing brain disorders.

8.1 Introduction

Depression is considered a significant public health concern, affecting individuals globally and presenting a considerable burden on both healthcare systems and society as a whole [156, 67]. Despite its prevalence, the accurate diagnosis of depression remains challenging, often relying on subjective assessments by clinicians based on criteria outlined in the Diagnostic and Statistical Manual of Mental Disorders [9]. However, the subjective nature of these assessments can lead to misdiagnosis and delays in appropriate treatment, highlight the critical need for more objective and reliable diagnostic methods [50].

Recent advancements in neuroimaging techniques, coupled with the advent of artificial intelligence (AI), offer promising avenues for enhancing the accuracy of depression diagnosis [120, 173]. Neuroimaging modalities, such as Fluorodeoxyglucose Positron Emission Tomography (FDG-PET) images, provide valuable insights into the functional alterations in the brain associated with depression [204]. By leveraging these neuroimaging data in conjunction with AI algorithms, we can potentially uncover patterns indicative of depression with greater sensitivity and specificity than traditional diagnostic approaches [71].

Among the various AI analytical tools available, Graph Neural Network (GNN) has emerged as a powerful framework for analyzing the brain from a network point of view. Graph-based representations capture the interactions between brain regions, enabling the characterization of network topology and identification of aberrant connectivity patterns associated with brain disorders [87, 200]. GNNs offer the capability to automatically learn features from graph-structured data, potentially capturing subtle but discriminative patterns relevant to brain disorders, and have recently attracted a lot of attention in the field of neuroscience [19]. While there are various studies on graph neural networks for brain disorders such as Alzheimer's Disease and Parkinson's Diagnosis [243], the applications of GNNs in FDG-PET analysis, as well as their utility in depression diagnosis, are very limited. Secondly, previous studies on graph and FDG-PET images mostly focused on pairwise graphs, where connections are established between only two nodes. However, given the brain is a complex system,

pairwise relationships may be insufficient to represent the complex relationships [242, 110]. Thirdly, compared to brain disorders such as Alzheimer's Disease, the differences in FDG-PET images of depression patients are generally more subtle. Therefore, we need a more powerful tool to capture this complexity.

In this work, we aim to fill a gap in the research by investigating the effectiveness of various GNN models and hypergraphs in analyzing FDG-PET images for diagnosing depression. Hypergraphs extend the capabilities of traditional graphs by allowing edges to connect any number of vertices, enabling a more adaptable representation of complex relationships. We hypothesize that utilizing hypergraph representations along with GNNs, which are powerful in learning intricate features from graph data, can enhance the accuracy and reliability of depression diagnosis compared to using traditional graphs. Our main contributions are summarized as follows

- We propose a pipeline for building a pairwise graph from static FDG-PET images using Kernel Density Estimation (KDE) and Dynamic Time Warping (DTW). Based on that, we construct a hypergraph to represent the higher information from data.
- We conduct experiments on a local dataset to compare the effectiveness of pairwise graphs and hypergraphs on several GNN classifiers, including Graph Convolutional Network and Graph Isomorphism Network. The experimental results show promising improvements in depression diagnosis with hypergraphs.

This article is structured as follows: Section 2 provides an overview of the data and methods used in our study, including the construction of graph and hypergraph representations, as well as the utilization of GNN models. Section 3 presents the results of our experiments and discusses the implications of our findings. Finally, Section 4 concludes the article by summarizing the contributions of our work and outlining directions for future research.

8.2 Materials and Methods

The flowchart of our study, as shown in Fig. 8.1, includes steps for building the graph (blocks in blue color) and performance evaluation using GNNs (blocks in red color). In the following subsections, we introduce each of these blocks in detail.

8.2.1 Datasets

The dataset used in our study comprises static FDG-PET images from 105 subjects, including 60 healthy individuals and 45 individuals diagnosed with depression. These images were gathered through clinical trials referenced as NCT00484523 and NCT02559466, corresponding to the healthy and depression groups, respectively.

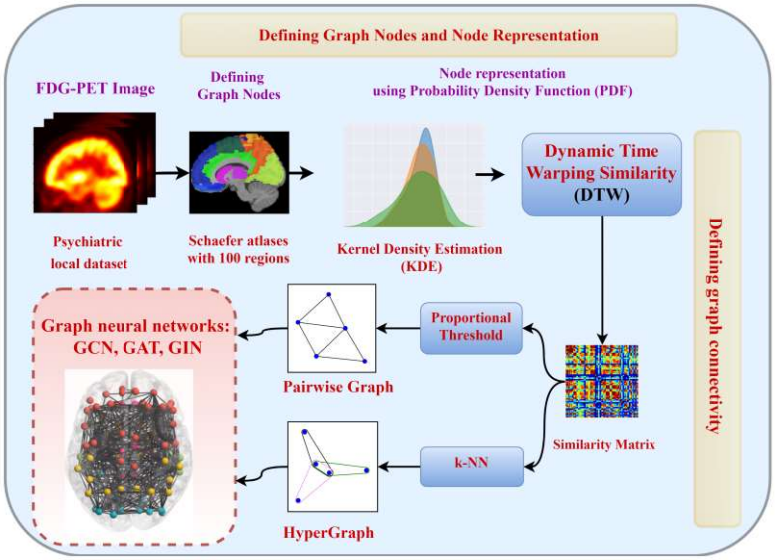

Figure 8.1: Pipeline of experiments.

8.2.2 Methods

8.2.2.1 Graph and Hypergraph Annotations

Graphs and hypergraphs are fundamental mathematical structures used to represent relationships between entities in various domains. A graph $G = (V, E)$ consists of a set of vertices V and a set of edges E, where each edge is a pair of vertices. Notationally, an edge $e \in E$ can be represented as $e = (v_i, v_j)$, indicating a connection between vertices v_i and v_j. In contrast, a hypergraph $H = (V, \mathscr{E})$ generalizes the notion of a graph by allowing edges to connect more than two vertices. Here, $\mathscr{E}$ denotes a set of hyperedges, each of which can connect an arbitrary subset of vertices. Mathematically, a hyperedge $e \in \mathscr{E}$ can be represented as $e = \{v_1, v_2, ..., v_k\}$, where $v_1, v_2, ..., v_k \in V$ are the vertices incident to the hyperedge. The use of graph and hypergraph notations provides a concise and expressive way to model complex relationships and dependencies among entities in various applications [74, 11].

8.2.2.2 Pairwise Graph and Hypergraph Building

Pairwise Graph Constructing a graph from a single static FDG-PET image poses a challenge. Unlike functional magnetic resonance imaging (fMRI) images in 4D (three spatial dimensions plus time), static FDG-PET images are only in 3D. Therefore, it is a bit difficult to construct brain connectivity to represent dynamic patterns of brain activity from FDG-PET images. To overcome this limitation, several methods based on KDE have been proposed to create individual brain networks from static FDG-PET images, showing advantages for Alzheimer's diagnosis [229, 213]. In this study, we follow the steps outlined in previous work [213] to construct a graph from FDG-PET images. Specifically, for each static FDG-PET image, we utilize the Schaefer

atlas [186], comprising 100 brain regions, to segment the brain images into regions of interest (ROIs). We treat these ROIs as nodes in the graph. Next, we represent the metabolic activity of each region as a Probability Density Function using KDE [233]. This approach ensures a smooth and accurate depiction of ROI characteristics. Subsequently, to establish connectivity between two nodes of the graph, we apply DTW [189] to quantify the similarity between two PDFs. Following these steps, from a single FDG-PET image, we obtain a connectivity matrix in which each element represents the similarity of PDFs between two nodes. Further details on the steps for building a graph from static FDG-PET images can be found in [213].

Hypergraph In constructing the hypergraph, the type of hyperedge determines the approach employed, which may include distance-based, representation-based, attributed-based, or network-based methods [74]. In this study, for simplicity, we utilize the similarity matrix derived from the pairwise graph and apply the k-NN method for constructing hyperedges. For each node in the graph, we create hyperedges connecting it to its k-nearest neighbors, based on the similarity of connectivity patterns between this node and others. This process involves identifying the k nodes with the closest connectivity profiles to each node in the graph and forming hyperedges to encapsulate these connections. Here, we experiment with different k-values in the k-NN method to determine the most suitable representation for hyperedges.

In our case, the relationship between graph and hypergraph construction methods lies in their representation of connectivity patterns and the incorporation of higher-order interactions. In graph construction, the similarity matrix captures pairwise similarity between nodes, representing connectivity patterns based on the similarity of metabolic activity profiles. Conversely, constructing the hypergraph involves forming each hyperedge based on the similarity between a node and its k-nearest neighbors, considering how this similarity interacts with all other nodes in the graph. This higher-order representation encapsulates the complex nature of connectivity within the brain network, capturing interactions beyond pairwise relationships. By incorporating these higher-order interactions, the hypergraph is expected to provide a more complex representation of brain connectivity, thereby enhancing our ability to discern subtle patterns relevant to depression diagnosis.

8.2.2.3 Graph Neural Networks

Graph Neural Networks have emerged as powerful tools for learning features from graph-structured data, garnering significant attention across various fields, including neuroscience [19]. Within the realm of GNNs, Graph Convolutional Networks (GCNs) [112] and Graph Isomorphism Networks (GINs) [235] stand out prominently. Given our focus on graph classification tasks in this work, we primarily concentrate on GCNs and GINs. GCNs are well-known in the GNN literature for their great performance in capturing graph structural information for various tasks, including graph classification. On the other hand, GINs are specifically tailored for classification tasks, making them a suitable candidate for comparison in our study. While GCNs operate by

iteratively aggregating information from neighboring nodes, GINs utilize learnable functions to ensure permutation invariance, facilitating effective classification in diverse graph structures. For these reasons, GCNs and GINs are the primary focus of our investigation. More details about GCNs and GINs can be found in the respective references [112, 235].

8.2.3 Experimental settings

8.2.3.1 Network structure

In this study, we primarily focus on comparing the effectiveness of hypergraphs for capturing complex patterns in the data, with the expectation that they can improve the performance of GNN classifiers. To achieve this, we employ the same network structure and experimental settings for both pairwise and hypergraph representations. For the GCN, our architecture comprises two GCN Convolutional hidden layers. The first layer transforms input features from a dimension of 100 nodes to 128 channels, followed by the second layer further processing these features to a dimension of 64 channels. Each hidden layer is followed by a Rectified Linear Unit (ReLU) activation function and dropout regularization before passing the features through the layers. On the other hand, the GIN architecture consists of two GIN Convolutional layers, each employing a sequential neural network with linear transformations, batch normalization, and ReLU activation functions. Both GCN and GIN utilize a $MeanMax$ readout function to aggregate node features across the graph, followed by a fully connected layer with Softmax to map features to a 2-dimensional output space.

8.2.3.2 Network training

When training the pairwise graph-based models, we employ the adjacency matrix directly as features. To investigate the influence of graph sparsity on model performance and facilitate training, we apply thresholding to the adjacency matrix at various proportional sparsity levels (5% to 50%). However, training the hypergraph model presents challenges due to the complexity of hyperedges. To simplify this process, we utilize clique expansion [3] to expand the hypergraph. This transformation converts each hyperedge into multiple edges between two nodes within the hyperedge, allowing us to treat and train the hypergraph as a conventional graph.

8.2.3.3 Performance Evaluation

In this study, we partitioned the dataset using a 3:1:1 split ratio for training, validation, and testing, respectively. To assess model performance, we utilized standard metrics, including balanced accuracy (ACC), sensitivity (SEN), and specificity (SPE). To mitigate bias, each experiment was repeated five times, and the average results were reported to confidently assess the effectiveness of each GNN model in diagnosing depression using FDG-PET images.

Table 8.1: Comparison of Performance Metrics for GNN models with pairwise graphs.

Sparse	**# Edges**	***GIN***			***GCN***		
		ACC (%)	SEN (%)	SPE (%)	ACC (%)	SEN (%)	SPE (%)
0.05	500	73.22	68.08	78.32	53.32	40.00	66.68
0.10	1000	68.46	56.96	79.98	54.44	8.88	100.00
0.15	1500	74.16	63.34	85.02	52.50	46.66	58.32
0.20	2000	64.86	56.38	73.34	72.78	65.56	80.02
0.25	2500	67.06	60.84	73.32	60.96	33.62	88.34
0.30	3000	63.22	48.04	78.34	59.04	34.74	83.34
0.35	3500	68.48	58.62	78.34	73.60	60.56	86.68
0.40	4000	67.10	65.84	68.32	68.78	70.86	66.68
0.45	4500	71.38	56.12	86.64	75.82	63.34	88.36
0.50	5000	62.64	50.28	75.02	71.12	57.24	84.98
Avg.* ± *std		68.06 ± 3.95	58.45 ± 6.36	77.66 ± 5.51	64.24 ± 9.15	48.15 ± 19.11	80.34 ± 12.67

Table 8.2: Comparison of Performance Metrics for GNN models with hypergraphs.

k	**# Edges**	***GIN***			***GCN***		
		ACC (%)	SEN (%)	SPE (%)	ACC (%)	SEN (%)	SPE (%)
3	272	64.56	45.82	83.34	59.16	40.00	78.32
5	507	71.68	61.68	81.68	66.94	42.24	91.66
7	734	65.84	55.00	76.68	66.24	35.84	96.66
10	1067	70.28	57.24	83.36	75.84	60.00	91.68
15	1626	69.18	55.00	83.34	71.80	50.28	93.34
20	2188	68.34	61.66	75.00	68.50	50.28	86.68
25	2753	72.80	58.90	86.68	78.90	66.14	91.66
30	3317	67.10	52.50	81.68	76.40	66.14	86.66
35	3881	63.34	41.68	85.00	78.32	68.36	88.34
40	4438	66.38	47.78	85.00	70.42	52.52	88.32
45	4988	72.34	59.74	84.98	71.94	57.22	86.68
50	5528	69.02	54.74	83.32	79.84	73.08	86.66
Avg.* ± *std		68.41 ± 3.06	54.31 ± 6.37	82.50 ± 3.44	72.02 ± 6.19	55.18 ± 12.02	88.89 ± 4.63

8.3 Results and Discussion

8.3.1 Results

Table 8.1 and 8.2 demonstrate that GNNs exhibit superior performance on hypergraphs compared to pairwise graphs. Specifically, the average accuracies of GIN and GCN models with hypergraphs are 68.41% and 72.02%, respectively, compared to 68.06% and 64.24% on pairwise graphs. Additionally, on average, the SPE of GIN and GCN improves by approximately 4.8% and 8.5%, respectively, on hypergraphs compared to conventional graphs. When comparing the performance of GCN and GIN, we observe that on conventional graphs, GIN performs better in terms of ACC and SEN. However, on hypergraphs, GCN outperforms GIN across all metrics. In the best-case scenario ($k = 35$), GCN improves performance by approximately 15%, 27%, and 3% compared to GIN. It reveals that GNNs demonstrate superior performance with hypergraphs compared to pairwise graphs. Specifically, the performance of GCN models on hypergraphs surpasses their performance on traditional graphs, indicating

the potential benefits of incorporating higher-order interactions in the analysis of brain connectivity networks.

Additionally, we observe that the sparsity of the network and the choice of k for constructing hyperedges significantly impact the overall performance of GNNs, both on graphs and hypergraphs. However, the trend is not as clear when the graph becomes sparser, emphasizing the importance of carefully selecting graph construction parameters. Furthermore, we also observe that the standard deviation of performance for GNN models on hypergraphs is lower, especially for GCN, compared to conventional graphs. For instance, the standard deviation of GCN for SPE on hypergraphs is 4.63, compared to 12.67 on graphs. Moreover, at the same number of edges, the performance of GNNs on hypergraphs is also better than on conventional graphs (Table 8.1, 8.2). This observation once again suggests the advantages of training GNNs with hypergraphs.

8.3.2 Discussion

Contribution of GNN for Depression diagnosis using FDG-PET images Our study delves into comparing the performance of GNN models for depression diagnosis using FDG-PET images on both conventional graphs and hypergraphs. Our results underscore the potential of hypergraphs and GNNs, particularly GCN, as powerful tools for unraveling the complexities of brain connectivity networks in mental health disorders. To our knowledge, there have been no studies on GCN for depression diagnosis using static FDG-PET images. However, in the case of fMRI, previous research has demonstrated GCN's capability in distinguishing individuals with major depressive disorder from healthy controls [169]. Directly comparing performance is challenging due to differences in neuroimaging modality and data size. Nonetheless, we observe that our performance with GCN on hypergraphs for FDG-PET is highly competitive with the performance of GCN on fMRI. This observation emphasizes the effectiveness of GCN for depression diagnosis while also highlighting the contribution of hypergraphs and GNNs to improving depression diagnosis using FDG-PET images.

Furthermore, by investing the impacts of graph network models (GIN, GCN), network sparsity, and graph representation (conventional, hypergraph) on model performance, our work offers valuable insights for researchers and practitioners regarding the roles and contributions of GNN models in diagnosing depression, particularly, and perhaps in the computational psychiatry community generally. This work is considered an initial step paving the way for more accurate and personalized diagnostic approaches.

Future Work While our experiments focused on a specific dataset and network architecture, further exploration with diverse datasets and alternative graph network architectures could enhance the robustness and generalizability of our findings. Additionally, exploring alternative methods for constructing graphs and hypergraphs from FDG-PET images, integrating networks at multiple scales, and fusion with other

modalities like structural MRI or functional MRI could provide valuable insights and enhance the diagnostic accuracy of depression.

8.4 Conclusion

In conclusion, our study underscores the significant contributions of Graph Neural Networks and hypergraphs in advancing depression diagnosis using FDG-PET images. Through the experimentation, we have demonstrated the superior performance of GNNs on hypergraphs compared to traditional graphs, emphasizing the potential of hypergraph representations in capturing complex neurobiological phenomena associated with depression. Furthermore, our exploration of GNNs also highlights the effectiveness of GCNs for depression diagnosis, showcasing the potential utility of GCNs and hypergraphs in deciphering intricate connectivity patterns within the brain. The integration of GNNs and hypergraph methodologies holds great promise for enhancing mental health diagnostics, thereby facilitating more accurate and personalized diagnostic approaches.

Chapter 9

Conclusion

Summary

This thesis is dedicated to addressing issues related to AD diagnosis using FDG-PET images and Machine Learning, focusing primarily on a region-based approach. It tackles and proposes methods to study several crucial aspects:

- **Selecting the most significant features affected during the course of AD:** Chapter 3 introduces a feature ranking method based on Autoencoder to select the most crucial features at various AD stages. This method is grounded in the concept of "simple-but-multiple" models and connection weights. It demonstrates the ability to select vital features while advocating for the use of multiple simple models instead of complex architectures.

- **The importance of brain atlas choices, the necessity for multiple-scale analysis, and the value of adapted brain mapping learned from data**: Chapter 4 presents a data-driven approach using FDG-PET images to construct an adaptive brain mapping for diagnosing AD. This chapter introduces the concept of a coefficient atlas, not only segmenting the brain for further studies but also quantifying the significance of each region in brain disorders. It emphasizes the requirement for multiple-scale analysis.

- **The role of graphs as tools for understanding interactions between multiple brain regions**: Recognizing the brain's complexity, Chapter 5 provides a concise overview of graph and metabolic graph analysis.

- **Individual brain networks for better comprehension and diagnosis of brain disorders**: Chapters 6 propose a method to construct individual brain networks. The results demonstrate the efficacy of these methods in capturing AD-related changes and individual variances.

- **Contribution of graph neural network for better diagnosis of brain disorders**: Chapter 7 presents a benchmark analysis of several graph neural network architectures and network building settings on the performance of classifiers for

both neurological disorders (AD) and psychological disorders (Depression). The results emphasize the contribution of graph neural networks as a framework for end-to-end learning and better disorder analysis. Chapter 8 investigates the effectiveness of hypergraphs as a tool for capturing the more complex patterns in patients' brain disorders, showing effectiveness especially in cases of psychological disorders, where the difference between patient and control groups is subtle.

Perspectives

Based on the work done, several considerations can be made for future research:

From a data perspective:

- **Multiple modalities:** It would be valuable to explore the application of the proposed methods on various types of neuroimaging and datasets to understand their advantages and disadvantages.
- **Cross-disorder analysis:** Studying the impact of different brain disorders on distinct regions is crucial. Different disorders may affect the same brain regions. The ability to differentiate between various types of cross-disorders is an essential task.

From a methodological standpoint:

- **Multiple scale analysis:** While this thesis underscores the importance of multiple scale analysis, detailed analysis remains limited. For future studies, exploring brain networks at various scales can uncover changes and precisely determine alterations at different stages of disorders. Furthermore, investigating methods that integrate information from different scales can enhance the diagnosis of disorders.
- **Variability between subjects:** With increasing interest in personalized medicine and improved treatments, understanding differences among subjects under various conditions—incorporating clinical and demographic information—is crucial for better diagnosis and treatment.
- **Graph neural networks:** Graphs serve as tools to study interactions among brain regions. Leveraging graph neural networks can significantly aid in exploring these intricate patterns extracted from graphs. Furthermore, GNNs possess the potential to not only enhance classifier performance but also to seamlessly integrate graph information across multiple neuroimaging modalities. However, there remains a need for further study in selecting optimal architecture, refining graph fusion techniques, and exploring additional architectures and methods for fusing graphs. Additionally, incorporating explainable deep learning approaches could provide valuable insights into the inner workings of these models, enhancing their interpretability and trustworthiness.

Clinical applications: All the work conducted thus far in this thesis has centered around classification perspectives. Translating these findings into clinical applications is crucial to assist doctors in making better decisions. For future research, delving into disease progression, exploring how ML contributes to improved treatments, investigating correlations between neuroimaging and clinical scores, and conducting comprehensive comparisons with other studies are all essential.

Bibliography

[ABZ19] Abubakar Abid et al. *Concrete Autoencoders for Differentiable Feature Selection and Reconstruction*. Jan. 2019. DOI: 10.48550/arXiv.1901.09346. arXiv: 1901.09346 [cs, stat]. (Visited on 03/27/2024) (cit. on pp. 54, 75).

[Ach+12] Radhakrishna Achanta et al. "SLIC Superpixels Compared to State-of-the-Art Superpixel Methods". In: *IEEE transactions on pattern analysis and machine intelligence* 34.11 (2012), pp. 2274–2282 (cit. on pp. 91, 94).

[ABB06] Sameer Agarwal et al. "Higher Order Learning with Graphs". In: *Proceedings of the 23rd International Conference on Machine Learning - ICML '06*. Pittsburgh, Pennsylvania: ACM Press, 2006, pp. 17–24. ISBN: 978-1-59593-383-6. DOI: 10.1145/1143844.1143847. (Visited on 03/13/2024) (cit. on p. 159).

[AAB16a] Ane Alberdi et al. "On the Early Diagnosis of Alzheimer's Disease from Multimodal Signals: A Survey". In: *Artificial Intelligence in Medicine* 71 (July 2016), pp. 1–29. ISSN: 09333657. DOI: 10.1016/j.artmed.2016.06.003 (cit. on p. 30).

[AAB16b] Ane Alberdi et al. "On the Early Diagnosis of Alzheimer's Disease from Multimodal Signals: A Survey". In: *Artificial intelligence in medicine* 71 (2016), pp. 1–29 (cit. on p. 90).

[AAB16c] Ane Alberdi et al. "On the Early Diagnosis of Alzheimer's Disease from Multimodal Signals: A Survey". In: *Artificial Intelligence in Medicine* 71 (July 2016), pp. 1–29. ISSN: 0933-3657. DOI: 10.1016/j.artmed.2016.06.003. (Visited on 03/27/2024) (cit. on p. 129).

[Alb+13] Marilyn S. Albert et al. "The Diagnosis of Mild Cognitive Impairment Due to Alzheimer's Disease: Recommendations from the National Institute on Aging-Alzheimer's Association Workgroups on Diagnostic Guidelines for Alzheimer's Disease". In: *FOCUS* 11.1 (Jan. 2013), pp. 96–106. ISSN: 1541-4094, 1541-4108. DOI: 10.1176/appi.focus.11.1.96 (cit. on pp. 20, 29).

[Alg+20] Sharifa Alghowinem et al. "Interpretation of Depression Detection Models via Feature Selection Methods". In: *IEEE Transactions on Affective Computing* 14.1 (Nov. 2020), pp. 133–152. ISSN: 1949-3045, 2371-9850. DOI: 10.1109/TAFFC.2020.3035535. (Visited on 03/27/2024) (cit. on p. 60).

[AA13] DSMTF American Psychiatric Association et al. *Diagnostic and Statistical Manual of Mental Disorders: DSM-5*. Vol. 5. American psychiatric association Washington, DC, 2013. (Visited on 03/13/2024) (cit. on p. 155).

[Amo+02] Piero Amodio et al. "Variability of Trail Making Test, Symbol Digit Test and Line Trait Test in Normal People. A Normative Study Taking into Account Age-Dependent Decline and Sociobiological Variables". In: *Aging Clinical and Experimental Research* 14.2 (Apr. 2002), pp. 117–131. ISSN: 1594-0667, 1720-8319. DOI: 10.1007/BF03324425 (cit. on p. 31).

[Ant+24] Alessia Antelmi et al. "A Survey on Hypergraph Representation Learning: ACM Computing Surveys". In: *ACM Computing Surveys* 56.1 (Jan. 2024), pp. 1–38. ISSN: 03600300. DOI: 10.1145/3605776. (Visited on 03/13/2024) (cit. on p. 157).

[Ars+18] Salim Arslan et al. "Human Brain Mapping: A Systematic Comparison of Parcellation Methods for the Human Cerebral Cortex". In: *Neuroimage* 170 (2018), pp. 5–30 (cit. on pp. 90, 91).

[AI02] David Attwell et al. "The Neural Basis of Functional Brain Imaging Signals". In: *Trends in Neurosciences* 25.12 (Dec. 2002), pp. 621–625. ISSN: 01662236. DOI: 10.1016/S0166-2236(02)02264-6. (Visited on 10/12/2023) (cit. on pp. 104, 117).

[Aze+09] Frederico A.C. Azevedo et al. "Equal Numbers of Neuronal and Nonneuronal Cells Make the Human Brain an Isometrically Scaled-up Primate Brain". In: *The Journal of Comparative Neurology* 513.5 (Apr. 2009), pp. 532–541. ISSN: 00219967, 10969861. DOI: 10.1002/cne.21974. (Visited on 10/12/2023) (cit. on p. 104).

[Bai+15] Matthieu Bailly et al. "Precuneus and Cingulate Cortex Atrophy and Hypometabolism in Patients with Alzheimer's Disease and Mild Cognitive Impairment: MRI and ^{18} F-FDG PET Quantitative Analysis Using FreeSurfer". In: *BioMed Research International* 2015 (2015), pp. 1–8. ISSN: 2314-6133, 2314-6141. DOI: 10.1155/2015/583931 (cit. on p. 35).

[BB06] Danielle Smith Bassett et al. "Small-World Brain Networks". In: *The Neuroscientist* 12.6 (Dec. 2006), pp. 512–523. ISSN: 1073-8584, 1089-4098. DOI: 10.1177/1073858406293182. (Visited on 10/12/2023) (cit. on p. 107).

[Bec+13] Georg A. Becker et al. "PET Quantification of ^{18} F-Florbetaben Binding to β-Amyloid Deposits in Human Brains". In: *Journal of Nuclear Medicine* 54.5 (May 2013), pp. 723–731. ISSN: 0161-5505, 2159-662X. DOI: 10.2967/jnumed.112.107185 (cit. on p. 35).

[BMR23a] Alaa Bessadok et al. "Graph Neural Networks in Network Neuroscience". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45.5 (May 2023), pp. 5833–5848. ISSN: 0162-8828, 2160-9292, 1939-3539. DOI: 10.1109/TPAMI.2022.3209686. (Visited on 10/12/2023) (cit. on p. 125).

[BMR23b] Alaa Bessadok et al. “Graph Neural Networks in Network Neuroscience”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45.5 (May 2023), pp. 5833–5848. ISSN: 1939-3539. DOI: 10.1109/TPAMI.2022.3209686. (Visited on 03/13/2024) (cit. on pp. 155, 158).

[] *Biomarkers In Risk Assessment: Validity And Validation (EHC 222, 2001)*. URL: https://www.inchem.org/documents/ehc/ehc/ehc222.htm#4.0 (visited on 11/20/2023) (cit. on p. 31).

[Bla+18] Ganna Blazhenets et al. “Principal Components Analysis of Brain Metabolism Predicts Development of Alzheimer Dementia”. In: *Journal of Nuclear Medicine* 60.6 (Nov. 2018), pp. 837–843. DOI: 10.2967/jnumed.118.219097 (cit. on p. 85).

[BZ09] Kaj Blennow et al. “Cerebrospinal Fluid Biomarkers for Alzheimer’s Disease”. In: *Journal of Alzheimer’s Disease* 18.2 (Aug. 26, 2009). Ed. by Garth Bissette, pp. 413–417. ISSN: 18758908, 13872877. DOI: 10.3233/JAD-2009-1177 (cit. on p. 38).

[Blu+13] Thomas Blumensath et al. “Spatially Constrained Hierarchical Parcellation of the Brain with Resting-State fMRI”. In: *Neuroimage* 76 (2013), pp. 313–324 (cit. on p. 91).

[Boh+11a] Nicolaas I. Bohnen et al. “Effectiveness and Safety of 18F-FDG PET in the Evaluation of Dementia: A Review of the Recent Literature”. In: *Journal of Nuclear Medicine* 53.1 (Dec. 2011), pp. 59–71. DOI: 10.2967/jnumed.111.096578 (cit. on p. 84).

[Boh+11b] Nicolaas I. Bohnen et al. “Effectiveness and Safety of 18F-FDG PET in the Evaluation of Dementia: A Review of the Recent Literature”. In: *Journal of Nuclear Medicine* 53.1 (Dec. 2011), pp. 59–71. DOI: 10.2967/jnumed.111.096578 (cit. on p. 100).

[Boh+12] Nicolaas I. Bohnen et al. “Effectiveness and Safety of ^{18} F-FDG PET in the Evaluation of Dementia: A Review of the Recent Literature”. In: *Journal of Nuclear Medicine* 53.1 (Jan. 2012), pp. 59–71. ISSN: 0161-5505, 2159-662X. DOI: 10.2967/jnumed.111.096578. (Visited on 03/27/2024) (cit. on p. 68).

[Boz02] M Bozzali. “White Matter Damage in Alzheimer’s Disease Assessed in Vivo Using Diffusion Tensor Magnetic Resonance Imaging”. In: *Journal of Neurology, Neurosurgery & Psychiatry* 72.6 (June 1, 2002), pp. 742–746. ISSN: 00223050. DOI: 10.1136/jnnp.72.6.742 (cit. on p. 34).

[Bra+99] Eva Braak et al. “Neuropathology of Alzheimer’s Disease: What Is New since A. Alzheimer?” In: *European Archives of Psychiatry and Clinical Neurosciences* 249.S3 (Dec. 15, 1999), S14–S22. ISSN: 0940-1334. DOI: 10.1007/PL00014168 (cit. on p. 28).

[Bre13] Alain Bretto. *Hypergraph Theory: An Introduction*. Mathematical Engineering. Heidelberg: Springer International Publishing, 2013. ISBN: 978-3-319-00079-4 978-3-319-00080-0. (Visited on 11/06/2023) (cit. on p. 124).

[Bri98] J.-P. Brion. "Neurofibrillary Tangles and Alzheimer's Disease". In: *European Neurology* 40.3 (1998), pp. 130–140. ISSN: 0014-3022, 1421-9913. DOI: 10.1159/000007969 (cit. on p. 28).

[BS09a] Ed Bullmore et al. "Complex Brain Networks: Graph Theoretical Analysis of Structural and Functional Systems". In: *Nature Reviews Neuroscience* 10.3 (Mar. 2009), pp. 186–198. ISSN: 1471-003X, 1471-0048. DOI: 10.1038/nrn2575. (Visited on 10/12/2023) (cit. on pp. 104, 112).

[BS09b] Ed Bullmore et al. "Complex Brain Networks: Graph Theoretical Analysis of Structural and Functional Systems". In: *Nature Reviews Neuroscience* 10.3 (Feb. 2009), pp. 186–198. DOI: 10.1038/nrn2575 (cit. on p. 133).

[BS09c] Ed Bullmore et al. "Complex Brain Networks: Graph Theoretical Analysis of Structural and Functional Systems". In: *Nature Reviews Neuroscience* 10.3 (Feb. 2009), pp. 186–198. DOI: 10.1038/nrn2575 (cit. on p. 138).

[BS12] Ed Bullmore et al. "The Economy of Brain Network Organization". In: *Nature Reviews Neuroscience* 13.5 (May 2012), pp. 336–349. ISSN: 1471-003X, 1471-0048. DOI: 10.1038/nrn3214. (Visited on 10/12/2023) (cit. on p. 107).

[BB11] Edward T. Bullmore et al. "Brain Graphs: Graphical Models of the Human Brain Connectome". In: *Annual Review of Clinical Psychology* 7.1 (Apr. 2011), pp. 113–140. ISSN: 1548-5943, 1548-5951. DOI: 10.1146/annurev-clinpsy-040510-143934. (Visited on 10/12/2023) (cit. on pp. 104, 106).

[Cab+15a] Carlos Cabral et al. "Predicting Conversion from MCI to AD with FDG-PET Brain Images at Different Prodromal Stages". In: *Computers in Biology and Medicine* 58 (Mar. 2015), pp. 101–109. ISSN: 0010-4825. DOI: 10.1016/j.compbiomed.2015.01.003. (Visited on 03/27/2024) (cit. on p. 61).

[Cab+15b] Carlos Cabral et al. "Predicting Conversion from MCI to AD with FDG-PET Brain Images at Different Prodromal Stages". In: *Computers in biology and medicine* 58 (2015), pp. 101–109 (cit. on p. 90).

[CF05] Marco Catani et al. "The Rises and Falls of Disconnection Syndromes". In: *Brain* 128.10 (Oct. 2005), pp. 2224–2239. ISSN: 1460-2156, 0006-8950. DOI: 10.1093/brain/awh622. (Visited on 10/12/2023) (cit. on p. 118).

[Cer+03] R. Ceravolo et al. "Dopaminergic Degeneration and Perfusional Impairment in Lewy Body Dementia and Alzheimer?S Disease". In: *Neurological Sciences* 24.3 (Oct. 1, 2003), pp. 162–163. ISSN: 1590-1874, 1590-3478. DOI: 10.1007/s10072-003-0110-6 (cit. on p. 37).

[Cha+09] R. Chaves et al. "SVM-based Computer-Aided Diagnosis of the Alzheimer's Disease Using *t*-Test NMSE Feature Selection with Feature Correlation Weighting". In: *Neuroscience Letters* 461.3 (Sept. 2009), pp. 293–297. ISSN: 0304-3940. DOI: 10.1016/j.neulet.2009.06.052. (Visited on 03/27/2024) (cit. on p. 61).

[CB12] Cynthia C. Chernecky et al. *Laboratory Tests and Diagnostic Procedures*. Elsevier Health Sciences, Dec. 1, 2012. 1235 pp. ISBN: 978-1-4557-4502-9 (cit. on p. 38).

[Cho+22] Kyoo Ho Cho et al. "Metabolic Network Is Related to Surgical Outcome in Temporal Lobe Epilepsy with Hippocampal Sclerosis: A Brain FDG-PET Study". In: *Journal of Neuroimaging* 32.2 (2022), pp. 300–313. ISSN: 1552-6569. DOI: 10.1111/jon.12941. (Visited on 10/05/2023) (cit. on pp. 118, 127).

[Cho+09] Seok Rye Choi et al. "Preclinical Properties of ^{18}F-AV-45: A PET Agent for Aβ Plaques in the Brain". In: *Journal of Nuclear Medicine* 50.11 (Nov. 2009), pp. 1887–1894. ISSN: 0161-5505, 2159-662X. DOI: 10.2967/jnumed.109.065284 (cit. on p. 35).

[Chu+08] Terence C Chua et al. "Diffusion Tensor Imaging in Mild Cognitive Impairment and Alzheimer's Disease: A Review:" in: *Current Opinion in Neurology* 21.1 (Feb. 2008), pp. 83–92. ISSN: 1350-7540. DOI: 10.1097/WCO.0b013e3282f4594b (cit. on p. 34).

[CPM87] Harry T. Chugani et al. "Positron Emission Tomography Study of Human Brain Functional Development". In: *Annals of Neurology* 22.4 (Oct. 1987), pp. 487–497. ISSN: 0364-5134, 1531-8249. DOI: 10.1002/ana.410220408. (Visited on 10/12/2023) (cit. on p. 106).

[Col+02] Sean J. Colloby et al. "A Comparison of 99mTc-HMPAO SPET Changes in Dementia with Lewy Bodies and Alzheimer's Disease Using Statistical Parametric Mapping". In: *European Journal of Nuclear Medicine and Molecular Imaging* 29.5 (May 2002), pp. 615–622. ISSN: 1619-7070, 1619-7089. DOI: 10.1007/s00259-002-0778-5 (cit. on p. 37).

[Cor+10] María M. Corrada et al. "Dementia Incidence Continues to Increase with Age in the Oldest Old: The 90+ Study". In: *Annals of Neurology* 67.1 (Jan. 2010), pp. 114–121. ISSN: 0364-5134, 1531-8249. DOI: 10.1002/ana.21915 (cit. on p. 27).

[CV95a] Corinna Cortes et al. "Support Vector Machine". In: *Machine learning* 20.3 (1995), pp. 273–297 (cit. on p. 79).

[CV95b] Corinna Cortes et al. "Support Vector Machine". In: *Machine learning* 20.3 (1995), pp. 273–297 (cit. on p. 91).

[CHP19] JM Coughlin et al. "Opportunities in Precision Psychiatry Using PET Neuroimaging in Psychosis". In: *NEUROBIOLOGY OF DISEASE* 131 (Nov. 2019). ISSN: 0969-9961. DOI: 10.1016/j.nbd.2019.03.016 (cit. on p. 155).

[Cra+12] R Cameron Craddock et al. "A Whole Brain fMRI Atlas Generated via Spatially Constrained Spectral Clustering". In: *Human brain mapping* 33.8 (2012), pp. 1914–1928 (cit. on p. 91).

[Dau10] Mak Adam Daulatzai. "Early Stages of Pathogenesis in Memory Impairment during Normal Senescence and Alzheimer's Disease". In: *Journal of Alzheimer's Disease* 20.2 (Apr. 1, 2010), pp. 355–367. ISSN: 18758908, 13872877. DOI: 10.3233/JAD-2010-1374 (cit. on p. 28).

[De +19] Siemon C. De Lange et al. "Shared Vulnerability for Connectome Alterations across Psychiatric and Neurological Brain Disorders". In: *Nature Human Behaviour* 3.9 (Aug. 2019), pp. 988–998. ISSN: 2397-3374. DOI: 10.1038/s41562-019-0659-6. (Visited on 10/12/2023) (cit. on p. 123).

[dG14] Juan de Oña et al. "Extracting the Contribution of Independent Variables in Neural Network Models: A New Approach to Handle Instability". In: *Neural Computing and Applications* 25.3 (Sept. 2014), pp. 859–869. ISSN: 1433-3058. DOI: 10.1007/s00521-014-1573-5. (Visited on 03/27/2024) (cit. on pp. 52, 55).

[Del+08] Angelo Del Sole et al. "Individual Cerebral Metabolic Deficits in Alzheimer's Disease and Amnestic Mild Cognitive Impairment: An FDG PET Study". In: *European Journal of Nuclear Medicine and Molecular Imaging* 35.7 (July 2008), pp. 1357–1366. ISSN: 1619-7070, 1619-7089. DOI: 10.1007/s00259-008-0773-6 (cit. on p. 35).

[Del+14] Julien Delrieu et al. "Apathy as a Feature of Prodromal Alzheimers Disease: An FDG-PET ADNI Study". In: *International Journal of Geriatric Psychiatry* 30.5 (June 2014), pp. 470–477. DOI: 10.1002/gps.4161 (cit. on p. 85).

[DO21] William O. Tatum DO IV. *Handbook of EEG Interpretation*. Springer Publishing Company, May 7, 2021. 423 pp. ISBN: 978-0-8261-4709-7. Google Books: 00oUEAAAQBAJ (cit. on p. 38).

[DS19] Guillaume Doquet et al. "Agnostic Feature Selection". In: *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer, 2019, pp. 343–358 (cit. on p. 75).

[DS20] Guillaume Doquet et al. "Agnostic Feature Selection". In: *Machine Learning and Knowledge Discovery in Databases*. Ed. by Ulf Brefeld et al. Vol. 11906. Cham: Springer International Publishing, 2020, pp. 343–358. ISBN: 978-3-030-46149-2 978-3-030-46150-8. DOI: 10.1007/978-3-030-46150-8_21. (Visited on 03/27/2024) (cit. on p. 54).

[Dua+11] Ranjan Duara et al. "Pre-MCI and MCI: Neuropsychological, Clinical, and Imaging Features and Progression Rates". In: *The American Journal of Geriatric Psychiatry* 19.11 (Nov. 2011), pp. 951–960. DOI: 10.1097/jgp.0b013e3182107c69 (cit. on p. 85).

[Dub+07] Bruno Dubois et al. "Research Criteria for the Diagnosis of Alzheimer's Disease: Revising the NINCDS–ADRDA Criteria". In: *The Lancet Neurology* 6.8 (Aug. 2007), pp. 734–746. ISSN: 14744422. DOI: 10.1016/S1474-4422(07)70178-3 (cit. on pp. 29, 30).

[Dub+14] Bruno Dubois et al. "Advancing Research Diagnostic Criteria for Alzheimer's Disease: The IWG-2 Criteria". In: *The Lancet Neurology* 13.6 (June 2014), pp. 614–629. ISSN: 14744422. DOI: 10.1016/S1474-4422(14)70090-0 (cit. on p. 36).

[Dub+16] Bruno Dubois et al. "Preclinical Alzheimer's Disease: Definition, Natural History, and Diagnostic Criteria". In: *Alzheimer's & Dementia* 12.3 (Mar. 2016), pp. 292–323. ISSN: 1552-5260, 1552-5279. DOI: 10.1016/j.jalz.2016.02.002 (cit. on pp. 20, 29).

[EYG18] Simon B. Eickhoff et al. "Imaging-Based Parcellations of the Human Brain". In: *Nature Reviews Neuroscience* 19.11 (Oct. 2018), pp. 672–686. DOI: 10.1038/s41583-018-0071-7 (cit. on p. 91).

[ETv22] Ahmed ElGazzar et al. *Benchmarking Graph Neural Networks for FMRI Analysis*. Nov. 2022. arXiv: 2211.08927 [cs]. (Visited on 03/13/2024) (cit. on pp. 140, 145).

[FKL19] Farzad V. Farahani et al. "Application of Graph Theory for Identifying Connectivity Patterns in Human Brain Networks: A Systematic Review". In: *Frontiers in Neuroscience* 13 (June 2019), p. 585. ISSN: 1662-453X. DOI: 10.3389/fnins.2019.00585. (Visited on 10/12/2023) (cit. on pp. 104, 123).

[Fer+22] AJ Ferrari et al. "Global, Regional, and National Burden of 12 Mental Disorders in 204 Countries and Territories, 1990-2019: A Systematic Analysis for the Global Burden of Disease Study 2019". In: *LANCET PSYCHIATRY* 9.2 (Feb. 2022), pp. 137–150. ISSN: 2215-0374. DOI: 10.1016/S2215-0366(21)00395-3 (cit. on p. 155).

[Fer+14] Daniel Ferreira et al. "Meta-Review of CSF Core Biomarkers in Alzheimerâ€™s Disease: The State-of-the-Art after the New Revised Diagnostic Criteria". In: *Frontiers in Aging Neuroscience* 6 (Mar. 24, 2014). ISSN: 1663-4365. DOI: 10.3389/fnagi.2014.00047 (cit. on p. 27).

[For+17] Stefania Forner et al. "Synaptic Impairment in Alzheimer's Disease: A Dysregulated Symphony". In: *Trends in Neurosciences* 40.6 (June 2017), pp. 347–357. ISSN: 01662236. DOI: 10.1016/j.tins.2017.04.002 (cit. on p. 19).

[FK99] H. Förstl et al. "Clinical Features of Alzheimer's Disease". In: *European Archives of Psychiatry and Clinical Neuroscience* 249.6 (Dec. 1, 1999), pp. 288–290. ISSN: 1433-8491. DOI: 10.1007/s004060050101 (cit. on p. 19).

[Fra+24] Francesco Fraioli et al. "Perspectives of the European Association of Nuclear Medicine on the Role of Artificial Intelligence (AI) in Molecular Brain Imaging". In: *European Journal of Nuclear Medicine and Molecular Imaging* 51.4 (Mar. 2024), pp. 1007–1011. ISSN: 1619-7089. DOI: 10.1007/s00259-023-06553-1. (Visited on 03/13/2024) (cit. on pp. 138, 139, 155).

[16] *Fundamentals of Brain Network Analysis*. Elsevier, 2016. ISBN: 978-0-12-407908-3. (Visited on 10/12/2023) (cit. on pp. 109, 112).

[Gal+12] E. Gallego-Jutgla et al. "Diagnosis of Alzheimer's Disease from EEG by Means of Synchrony Measures in Optimized Frequency Bands". In: *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 2012 34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). San Diego, CA: IEEE, Aug. 2012, pp. 4266–4270. ISBN: 978-1-4577-1787-1 978-1-4244-4119-8. DOI: 10.1109/EMBC.2012.6346909 (cit. on pp. 19, 27).

[Gao+22] Yue Gao et al. "Hypergraph Learning: Methods and Practices". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44.5 (May 2022), pp. 2548–2566. ISSN: 1939-3539. DOI: 10.1109/TPAMI.2020.3039374. (Visited on 03/13/2024) (cit. on pp. 157, 158).

[Gar+18] Imène Garali et al. "Histogram-Based Features Selection and Volume of Interest Ranking for Brain PET Image Classification". In: *IEEE Journal of Translational Engineering in Health and Medicine* 6 (2018), pp. 1–12. ISSN: 2168-2372. DOI: 10.1109/JTEHM.2018.2796600. (Visited on 03/27/2024) (cit. on p. 61).

[Ge+20] Jingjie Ge et al. "Metabolic Network as an Objective Biomarker in Monitoring Deep Brain Stimulation for Parkinson's Disease: A Longitudinal Study". In: *EJNMMI Research* 10.1 (Oct. 2020), p. 131. ISSN: 2191-219X. DOI: 10.1186/s13550-020-00722-1. (Visited on 10/06/2023) (cit. on pp. 119, 127).

[GDL03] Muriel Gevrey et al. "Review and Comparison of Methods to Study the Contribution of Variables in Artificial Neural Network Models". In: *Ecological Modelling*. Modelling the Structure of Acquatic Communities: Concepts, Methods and Problems. 160.3 (Feb. 2003), pp. 249–264. ISSN: 0304-3800. DOI: 10.1016/S0304-3800(02)00257-0. (Visited on 03/27/2024) (cit. on p. 52).

[Gra+12] Katherine R Gray et al. "Multi-Region Analysis of Longitudinal FDG-PET for the Classification of Alzheimer's Disease". In: *NeuroImage* 60.1 (2012), pp. 221–229 (cit. on pp. 90, 95).

[Gre+04] Michael D. Greicius et al. "Default-Mode Network Activity Distinguishes Alzheimer's Disease from Healthy Aging: Evidence from Functional MRI". In: *Proceedings of the National Academy of Sciences* 101.13 (Mar. 30, 2004), pp. 4637–4642. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.0308627101 (cit. on p. 33).

[Gue+06] E. Guedj et al. "Identification of Subgroups in Amnestic Mild Cognitive Impairment". In: *Neurology* 67.2 (July 25, 2006), pp. 356–358. ISSN: 0028-3878, 1526-632X. DOI: 10.1212/01.wnl.0000225076.73312.d4. pmid: 16864841 (cit. on p. 37).

[Gue+22] Eric Guedj et al. "EANM Procedure Guidelines for Brain PET Imaging Using [18F]FDG, Version 3". In: *European Journal of Nuclear Medicine and Molecular Imaging* 49.2 (Jan. 2022), pp. 632–651. ISSN: 1619-7089. DOI: 10.1007/s00259-021-05603-w. (Visited on 03/13/2024) (cit. on pp. 138, 139).

[Gun+22] Serafettin Gunes et al. "Biomarkers for Alzheimer's Disease in the Current State: A Narrative Review". In: *International Journal of Molecular Sciences* 23.9 (Apr. 29, 2022), p. 4962. ISSN: 1422-0067. DOI: 10.3390/ijms23094962 (cit. on p. 31).

[Guy+02] Isabelle Guyon et al. "Gene Selection for Cancer Classification Using Support Vector Machines". In: *Machine learning* 46.1-3 (2002), pp. 389–422 (cit. on p. 94).

[Hag+08] Patric Hagmann et al. "Mapping the Structural Core of Human Cerebral Cortex". In: *PLoS Biology* 6.7 (July 2008). Ed. by Karl J Friston, e159. DOI: 10.1371/journal.pbio.0060159 (cit. on p. 91).

[Han+18] Kai Han et al. "Autoencoder Inspired Unsupervised Feature Selection". In: *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Apr. 2018, pp. 2941–2945. DOI: 10.1109/ICASSP.2018.8462261. (Visited on 03/27/2024) (cit. on pp. 54, 75).

[Han+22] Oskar Hansson et al. "The Alzheimer's Association Appropriate Use Recommendations for Blood Biomarkers in Alzheimer's Disease". In: *Alzheimer's & Dementia* 18.12 (Dec. 2022), pp. 2669–2686. ISSN: 1552-5260, 1552-5279. DOI: 10.1002/alz.12756 (cit. on p. 38).

[HE10] Yong He et al. "Graph Theoretical Modeling of Brain Connectivity". In: *Current Opinion in Neurology* 23.4 (Aug. 2010), pp. 341–350. ISSN: 1350-7540. DOI: 10.1097/WCO.0b013e32833aa567. (Visited on 10/12/2023) (cit. on p. 155).

[Heb+13] Liesi E. Hebert et al. “Alzheimer Disease in the United States (2010–2050) Estimated Using the 2010 Census”. In: *Neurology* 80.19 (May 7, 2013), pp. 1778–1783. ISSN: 0028-3878, 1526-632X. DOI: 10.1212/WNL.0b013e31828726f5 (cit. on p. 19).

[Hel+08] E. P. Helzner et al. “Survival in Alzheimer Disease: A Multiethnic, Population-Based Study of Incident Cases”. In: *Neurology* 71.19 (Nov. 4, 2008), pp. 1489–1495. ISSN: 0028-3878, 1526-632X. DOI: 10.1212/01.wnl.0000334278.11022.42 (cit. on p. 30).

[Hin+09] Chris Hinrichs et al. “Spatially Augmented LPboosting for AD Classification with Evaluations on the ADNI Dataset”. In: *Neuroimage* 48.1 (2009), pp. 138–149 (cit. on p. 90).

[HZ93] Geoffrey E Hinton et al. “Autoencoders, Minimum Description Length and Helmholtz Free Energy”. In: *Advances in Neural Information Processing Systems*. Vol. 6. Morgan-Kaufmann, 1993. (Visited on 03/27/2024) (cit. on p. 57).

[HZ94] Geoffrey E. Hinton et al. “Autoencoders, Minimum Description Length and Helmholtz Free Energy”. In: *Advances in Neural Information Processing Systems*. Advances in neural information processing systems, 1994, pp. 3–10 (cit. on p. 77).

[Hof+00] John M. Hoffman et al. “FDG PET Imaging in Patients with Pathologically Verified Dementia”. In: *Journal of Nuclear Medicine* 41.11 (2000), pp. 1920–1928. ISSN: 0161-5505. eprint: https://jnm.snmjournals.org/content/41/11/1920.full.pdf (cit. on pp. 84, 85).

[Hu+15] Yuxiao Hu et al. “Small-Worldness and Gender Differences of Large Scale Brain Metabolic Covariance Networks in Young Adults: A FDG PET Study of 400 Subjects”. In: *Acta Radiologica* 56.2 (Feb. 2015), pp. 204–213. ISSN: 0284-1851, 1600-0455. DOI: 10.1177/0284185114529106. (Visited on 10/05/2023) (cit. on pp. 113, 117, 124, 127).

[Hua+22a] Gan Huang et al. “Cerebral Metabolic Network in Patients With Anti-N-Methyl-D-Aspartate Receptor Encephalitis on 18F-FDG PET Imaging”. In: *Frontiers in Neuroscience* 16 (2022). ISSN: 1662-453X. (Visited on 10/06/2023) (cit. on pp. 122, 127).

[HA07] Juebin Huang et al. “Diffusion Tensor Imaging of Normal Appearing White Matter and Its Correlation with Cognitive Functioning in Mild Cognitive Impairment and Alzheimer’s Disease”. In: *Annals of the New York Academy of Sciences* 1097.1 (Feb. 2007), pp. 259–264. ISSN: 0077-8923, 1749-6632. DOI: 10.1196/annals.1379.021 (cit. on p. 34).

[Hua+20a] Qi Huang et al. “Age-Associated Reorganization of Metabolic Brain Connectivity in Chinese Children”. In: *European Journal of Nuclear Medicine and Molecular Imaging* 47.2 (Feb. 2020), pp. 235–246. ISSN: 1619-7089. DOI: 10.1007/s00259-019-04508-z (cit. on pp. 113, 117, 124, 127).

[Hua+22b] Qi Huang et al. "Aging-Related Modular Architectural Reorganization of the Metabolic Brain Network". In: *Brain Connectivity* 12.5 (June 2022), pp. 432–442. ISSN: 2158-0014, 2158-0022. DOI: 10.1089/brain.2021.0054. (Visited on 10/11/2023) (cit. on pp. 113, 117, 127).

[Hua+18a] Sheng-Yao Huang et al. "Characteristic Patterns of Inter- and Intra-Hemispheric Metabolic Connectivity in Patients with Stable and Progressive Mild Cognitive Impairment and Alzheimer's Disease". In: *Scientific Reports* 8.1 (Sept. 2018), p. 13807. ISSN: 2045-2322. DOI: 10.1038/s41598-018-31794-8. (Visited on 10/06/2023) (cit. on pp. 113, 120, 121, 124, 127).

[Hua+18b] Sheng-Yao Huang et al. "Characteristic Patterns of Inter- and Intra-Hemispheric Metabolic Connectivity in Patients with Stable and Progressive Mild Cognitive Impairment and Alzheimer's Disease". In: *Scientific Reports* 8.1 (Sept. 2018), p. 13807. ISSN: 2045-2322. DOI: 10.1038/s41598-018-31794-8. (Visited on 10/06/2023) (cit. on p. 129).

[Hua+20b] Sheng-Yao Huang et al. "A Novel Individual Metabolic Brain Network for 18F-FDG PET Imaging". In: *Frontiers in Neuroscience* 14 (2020). ISSN: 1662-453X. (Visited on 10/06/2023) (cit. on pp. 113, 120, 121, 127).

[Hua+20c] Sheng-Yao Huang et al. "A Novel Individual Metabolic Brain Network for 18F-FDG PET Imaging". In: *Frontiers in Neuroscience* 14 (May 2020). DOI: 10.3389/fnins.2020.00344 (cit. on p. 129).

[Ima+20] Masamichi Imai et al. "Metabolic Network Topology of Alzheimer's Disease and Dementia with Lewy Bodies Generated Using Fluorodeoxyglucose Positron Emission Tomography". In: *Journal of Alzheimer's Disease* 73.1 (Jan. 2020), pp. 197–207. ISSN: 13872877, 18758908. DOI: 10.3233/JAD-190843. (Visited on 10/11/2023) (cit. on pp. 113, 122, 127).

[Irv16] Benjamin Irving. "maskSLIC: Regional Superpixel Generation with Application to Local Pathology Characterisation in Medical Images". In: *arXiv preprint arXiv:1606.09518* (2016). arXiv: 1606.09518 (cit. on p. 94).

[Jac+18] Clifford R. Jack et al. "NIA-AA Research Framework: Toward a Biological Definition of Alzheimer's Disease". In: *Alzheimer's & Dementia* 14.4 (Apr. 2018), pp. 535–562. ISSN: 1552-5260, 1552-5279. DOI: 10.1016/j.jalz.2018.02.018 (cit. on p. 31).

[Jag+01] W. Jagust et al. "SPECT Perfusion Imaging in the Diagnosis of Alzheimer's Disease: A Clinical-Pathologic Study". In: *Neurology* 56.7 (Apr. 10, 2001), pp. 950–956. ISSN: 0028-3878, 1526-632X. DOI: 10.1212/WNL.56.7.950 (cit. on p. 37).

[Jan+19] Iris E. Jansen et al. "Genome-Wide Meta-Analysis Identifies New Loci and Functional Pathways Influencing Alzheimer's Disease Risk". In: *Nature Genetics* 51.3 (Mar. 2019), pp. 404–413. ISSN: 1061-4036, 1546-1718. DOI: 10.1038/s41588-018-0311-9 (cit. on p. 38).

[Jes+14] Frank Jessen et al. "A Conceptual Framework for Research on Subjective Cognitive Decline in Preclinical Alzheimer's Disease". In: *Alzheimer's & Dementia* 10.6 (Nov. 2014), pp. 844–852. ISSN: 1552-5260, 1552-5279. DOI: 10.1016/j.jalz.2014.01.001 (cit. on pp. 19, 20, 29).

[Jie+16a] Biao Jie et al. "Hyper-Connectivity of Functional Networks for Brain Disease Diagnosis". In: *Medical Image Analysis* 32 (Aug. 2016), pp. 84–100. ISSN: 13618415. DOI: 10.1016/j.media.2016.03.003. (Visited on 11/06/2023) (cit. on p. 124).

[Jie+16b] Biao Jie et al. "Hyper-Connectivity of Functional Networks for Brain Disease Diagnosis". In: *Medical Image Analysis* 32 (Aug. 2016), pp. 84–100. ISSN: 1361-8423. DOI: 10.1016/j.media.2016.03.003 (cit. on p. 156).

[KG19] Rita Khoury et al. "Diagnostic biomarkers of Alzheimer's disease: a state-of-the-art review". In: *Biomarkers in Neuropsychiatry* 1 (2019), p. 100005 (cit. on p. 31).

[KW17] Thomas N. Kipf et al. *Semi-Supervised Classification with Graph Convolutional Networks*. Feb. 2017. arXiv: 1609.02907 [cs, stat]. (Visited on 03/13/2024) (cit. on pp. 158, 159).

[Klu98] W.E. Klunk. "Biological Markers of Alzheimer's Disease". In: *Neurobiology of Aging* 19.2 (Mar. 1998), pp. 145–147. ISSN: 01974580. DOI: 10.1016/S0197-4580(98)00013-X (cit. on p. 31).

[Klu+04] William E. Klunk et al. "Imaging Brain Amyloid in Alzheimer's Disease with Pittsburgh Compound-B". In: *Annals of Neurology* 55.3 (Mar. 2004), pp. 306–319. ISSN: 0364-5134, 1531-8249. DOI: 10.1002/ana.20009 (cit. on p. 35).

[Kov+18] Igor Koval et al. "Spatiotemporal Propagation of the Cortical Atrophy: Population and Individual Patterns". In: *Frontiers in Neurology* 9 (May 4, 2018), p. 235. ISSN: 1664-2295. DOI: 10.3389/fneur.2018.00235 (cit. on p. 32).

[Kua+19] Liqun Kuang et al. "Metabolic Brain Network Analysis of FDG-PET in Alzheimer's Disease Using Kernel-Based Persistent Features". In: *Molecules* 24.12 (Jan. 2019), p. 2301. ISSN: 1420-3049. DOI: 10.3390/molecules24122301. (Visited on 10/06/2023) (cit. on pp. 120, 124, 127).

[Lan+00] Jack L Lancaster et al. "Automated Talairach Atlas Labels for Functional Brain Mapping". In: *Human brain mapping* 10.3 (2000), pp. 120–131 (cit. on p. 90).

[Lee+14] Hyekyoung Lee et al. "Hole Detection in Metabolic Connectivity of Alzheimer's Disease Using k -Laplacian". In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2014*. Ed. by Polina Golland et al. Vol. 8675. Cham: Springer International Publishing, 2014, pp. 297–304. ISBN: 978-3-319-10442-3 978-3-319-10443-0. DOI: 10.1007/978-3-319-10443-0_38. (Visited on 10/06/2023) (cit. on pp. 113, 120, 127).

[Leu+22] Antoine Leuzy et al. "Blood-based Biomarkers for Alzheimer's Disease". In: *EMBO Molecular Medicine* 14.1 (Jan. 11, 2022), e14408. ISSN: 1757-4676, 1757-4684. DOI: 10.15252/emmm.202114408 (cit. on p. 38).

[Li+21] Fei Li et al. "Artificial Intelligence Applications in Psychoradiology". In: *Psychoradiology* 1.2 (June 2021), pp. 94–107. ISSN: 2634-4416. DOI: 10.1093/psyrad/kkab009. (Visited on 03/13/2024) (cit. on p. 155).

[Li+18a] Jundong Li et al. "Feature Selection: A Data Perspective". In: *ACM Computing Surveys* 50.6 (Nov. 2018), pp. 1–45. ISSN: 0360-0300, 1557-7341. DOI: 10.1145/3136625. (Visited on 03/27/2024) (cit. on pp. 53, 54, 60, 79).

[Li+23a] Yu-Lin Li et al. "Cross-Modality Comparison between Structural and Metabolic Networks in Individual Brain Based on the Jensen-Shannon Divergence Method: A Healthy Chinese Population Study". In: *Brain Structure & Function* 228.3-4 (May 2023), pp. 761–773. ISSN: 1863-2661. DOI: 10.1007/s00429-023-02616-z (cit. on pp. 104, 113, 117, 123, 124, 127).

[Li+23b] Weikai Li et al. "The Reconfiguration Pattern of Individual Brain Metabolic Connectome for Parkinson's Disease Identification". In: *MedComm* 4.4 (2023), e305. ISSN: 2688-2663. DOI: 10.1002/mco2.305. (Visited on 10/06/2023) (cit. on pp. 113, 119, 121, 124, 127).

[Li+18b] Ying Li et al. "Indirect Relation Based Individual Metabolic Network for Identification of Mild Cognitive Impairment". In: *Journal of Neuroscience Methods* 309 (Nov. 2018), pp. 188–198. ISSN: 01650270. DOI: 10.1016/j.jneumeth.2018.09.007. (Visited on 10/06/2023) (cit. on pp. 113, 120, 127).

[Lib+96] David J. Libon et al. "A Nine—Word Dementia Version of the California Verbal Learning Test". In: *The Clinical Neuropsychologist* 10.3 (July 1996), pp. 237–244. ISSN: 1385-4046, 1744-4144. DOI: 10.1080/13854049608406686 (cit. on p. 20).

[Liu+21] Jiayu Liu et al. "Regional Metabolic and Network Changes in Meige Syndrome". In: *Scientific Reports* 11.1 (Aug. 2021), p. 15753. ISSN: 2045-2322. DOI: 10.1038/s41598-021-95333-8. (Visited on 10/06/2023) (cit. on pp. 121, 127).

[Liu+22] Li Liu et al. "Reconfigured Metabolism Brain Network in Asymptomatic Microtubule-Associated Protein Tau Mutation Carriers: A Graph Theoretical Analysis". In: *Alzheimer's Research & Therapy* 14.1 (Apr. 2022), p. 52. ISSN: 1758-9193. DOI: 10.1186/s13195-022-01000-z. (Visited on 10/06/2023) (cit. on pp. 121, 127).

[Liu+12] Zhenyu Liu et al. "Altered Topological Patterns of Brain Networks in Mild Cognitive Impairment and Alzheimers Disease: A Resting-State fMRI Study". In: *Psychiatry Research: Neuroimaging* 202.2 (May 2012), pp. 118–125. DOI: 10.1016/j.pscychresns.2012.03.002 (cit. on pp. 134, 135).

[Lor+17] Louis-David Lord et al. "Understanding Principles of Integration and Segregation Using Whole-Brain Computational Connectomics: Implications for Neuropsychiatric Disorders". In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 375.2096 (June 2017), p. 20160283. ISSN: 1364-503X, 1471-2962. DOI: 10.1098/rsta.2016.0283. (Visited on 11/06/2023) (cit. on p. 107).

[Lyn+06] C. A. Lynch et al. "The Clinical Dementia Rating Sum of Box Score in Mild Dementia". In: *Dementia and Geriatric Cognitive Disorders* 21.1 (2006), pp. 40–43. ISSN: 1420-8008. DOI: 10.1159/000089218. pmid: 16254429 (cit. on p. 20).

[Ma+22] Jie Ma et al. "Brain Metabolic Network Redistribution in Patients with White Matter Hyperintensities on MRI Analyzed with an Individualized Index Derived from 18F-FDG-PET/MRI". In: *Korean Journal of Radiology* 23.10 (Oct. 2022), pp. 986–997. ISSN: 1229-6929. DOI: 10.3348/kjr.2022.0320. (Visited on 10/06/2023) (cit. on pp. 121, 127).

[MWH06] Robert W. Mahley et al. "Apolipoprotein E4: A Causative Factor and Therapeutic Target in Neuropathology, Including Alzheimer's Disease". In: *Proceedings of the National Academy of Sciences* 103.15 (Apr. 11, 2006), pp. 5644–5651. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.0600549103 (cit. on p. 38).

[MY19] Freddie Márquez et al. "Neuroimaging biomarkers for Alzheimer's disease". In: *Molecular neurodegeneration* 14 (2019), pp. 1–14 (cit. on p. 31).

[Mat09] Niklas Mattsson. "CSF Biomarkers and Incipient Alzheimer Disease in Patients With Mild Cognitive Impairment". In: *JAMA* 302.4 (July 22, 2009), p. 385. ISSN: 0098-7484. DOI: 10.1001/jama.2009.1064 (cit. on p. 29).

[McK+84] Guy McKhann et al. "Clinical Diagnosis of Alzheimer's Disease: Report of the NINCDS-ADRDA Work Group* under the Auspices of Department of Health and Human Services Task Force on Alzheimer's Disease". In: *Neurology* 34.7 (July 1984), pp. 939–939. ISSN: 0028-3878, 1526-632X. DOI: 10.1212/WNL.34.7.939 (cit. on p. 31).

[McK+11] Guy M. McKhann et al. "The Diagnosis of Dementia Due to Alzheimer's Disease: Recommendations from the National Institute on Aging-Alzheimer's Association Workgroups on Diagnostic Guidelines for Alzheimer's Disease". In: *Alzheimer's & Dementia* 7.3 (May 2011), pp. 263–269. ISSN: 1552-5260, 1552-5279. DOI: 10.1016/j.jalz.2011.03.005 (cit. on pp. 20, 31).

[MLB15] John D. Medaglia et al. "Cognitive Network Neuroscience". In: *Journal of Cognitive Neuroscience* 27.8 (Aug. 2015), pp. 1471–1491. ISSN: 0898-929X, 1530-8898. DOI: 10.1162/jocn_a_00810. (Visited on 10/12/2023) (cit. on pp. 104, 106).

[MLB10] David Meunier et al. "Modular and Hierarchically Modular Organization of Brain Networks". In: *Frontiers in Neuroscience* 4 (2010). ISSN: 1662-4548. DOI: 10.3389/fnins.2010.00200. (Visited on 10/12/2023) (cit. on p. 107).

[Mie+09] M.M. Mielke et al. "Regionally-Specific Diffusion Tensor Imaging in Mild Cognitive Impairment and Alzheimer's Disease". In: *NeuroImage* 46.1 (May 15, 2009), pp. 47–55. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2009.01.054 (cit. on p. 34).

[Mit+14] A. J. Mitchell et al. "Risk of Dementia and Mild Cognitive Impairment in Older People with Subjective Memory Complaints: Meta-Analysis". In: *Acta Psychiatrica Scandinavica* 130.6 (Dec. 2014), pp. 439–451. ISSN: 0001690X. DOI: 10.1111/acps.12336 (cit. on p. 20).

[Mit09] Alex J. Mitchell. "A Meta-Analysis of the Accuracy of the Mini-Mental State Examination in the Detection of Dementia and Mild Cognitive Impairment". In: *Journal of Psychiatric Research* 43.4 (Jan. 2009), pp. 411–431. ISSN: 00223956. DOI: 10.1016/j.jpsychires.2008.04.014 (cit. on pp. 20, 31).

[Moh+09] David Moher et al. "Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement". In: *Journal of Clinical Epidemiology* 62.10 (Oct. 2009), pp. 1006–1012. ISSN: 08954356. DOI: 10.1016/j.jclinepi.2009.06.005. (Visited on 10/12/2023) (cit. on p. 112).

[Mur+11] Hideo Mure et al. "Parkinson's Disease Tremor-Related Metabolic Network: Characterization, Progression, and Treatment Effects". In: *NeuroImage* 54.2 (Jan. 2011), pp. 1244–1253. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2010.09.028. (Visited on 10/06/2023) (cit. on pp. 119, 123, 127).

[Nas+05] Ziad S. Nasreddine et al. "The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment". In: *Journal of the American Geriatrics Society* 53.4 (Apr. 2005), pp. 695–699. ISSN: 0002-

8614, 1532-5415. DOI: 10.1111/j.1532-5415.2005.53221.x (cit. on p. 31).

[Nir+13] Talia M. Nir et al. "Effectiveness of Regional DTI Measures in Distinguishing Alzheimer's Disease, MCI, and Normal Aging". In: *NeuroImage: Clinical* 3 (2013), pp. 180–195. ISSN: 22131582. DOI: 10.1016/j.nicl.2013.07.006 (cit. on p. 34).

[Nor08] Agneta Nordberg. "Amyloid Plaque Imaging in Vivo: Current Achievement and Future Prospects". In: *European Journal of Nuclear Medicine and Molecular Imaging* 35.S1 (Mar. 2008), pp. 46–50. ISSN: 1619-7070, 1619-7089. DOI: 10.1007/s00259-007-0700-2 (cit. on p. 28).

[OJ02] Julian D Olden et al. "Illuminating the "Black Box": A Randomization Approach for Understanding Variable Contributions in Artificial Neural Networks". In: *Ecological Modelling* 154.1 (Aug. 2002), pp. 135–150. ISSN: 0304-3800. DOI: 10.1016/S0304-3800(02)00064-9. (Visited on 03/27/2024) (cit. on p. 55).

[OJD04] Julian D Olden et al. "An Accurate Comparison of Methods for Quantifying Variable Importance in Artificial Neural Networks Using Simulated Data". In: *Ecological Modelling* 178.3 (Nov. 2004), pp. 389–397. ISSN: 0304-3800. DOI: 10.1016/j.ecolmodel.2004.03.013. (Visited on 03/27/2024) (cit. on p. 52).

[Ota+15] Kenichi Ota et al. "Effects of Imaging Modalities, Brain Atlases and Feature Selection on Prediction of Alzheimer's Disease". In: *Journal of neuroscience methods* 256 (2015), pp. 168–183 (cit. on pp. 90, 95).

[Pad+12] P. Padilla et al. "NMF-SVM Based CAD Tool Applied to Functional Brain Images for the Diagnosis of Alzheimer's Disease". In: *IEEE Transactions on Medical Imaging* 31.2 (Feb. 2012), pp. 207–216. ISSN: 0278-0062, 1558-254X. DOI: 10.1109/TMI.2011.2167628. (Visited on 03/27/2024) (cit. on p. 61).

[Pag+15] M. Pagani et al. "Volume of Interest-Based [18F]Fluorodeoxyglucose PET Discriminates MCI Converting to Alzheimer's Disease from Healthy Controls. A European Alzheimer's Disease Consortium (EADC) Study". In: *NeuroImage: Clinical* 7 (Jan. 2015), pp. 34–42. ISSN: 2213-1582. DOI: 10.1016/j.nicl.2014.11.007. (Visited on 03/27/2024) (cit. on p. 61).

[Pag+17] Marco Pagani et al. "Early Identification of MCI Converting to AD: A FDG PET Study". In: *European Journal of Nuclear Medicine and Molecular Imaging* 44.12 (June 2017), pp. 2042–2052. DOI: 10.1007/s00259-017-3761-x (cit. on p. 85).

[Pak03] B Pakkenberg. "Aging and the Human Neocortex". In: *Experimental Gerontology* 38.1-2 (Jan. 2003), pp. 95–99. ISSN: 05315565. DOI: 10.1016/S0531-5565(02)00151-1. (Visited on 10/12/2023) (cit. on p. 104).

[Pan+18] Xiaoxi Pan et al. "Multilevel Feature Representation of FDG-PET Brain Images for Diagnosing Alzheimer's Disease". In: *IEEE journal of biomedical and health informatics* 23.4 (2018), pp. 1499–1506 (cit. on p. 84).

[PF13] Hae-Jeong Park et al. "Structural and Functional Brain Networks: From Connections to Cognition". In: *Science* 342.6158 (Nov. 2013), p. 1238411. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1238411. (Visited on 10/12/2023) (cit. on p. 104).

[Pat+16] V Patel et al. "Addressing the Burden of Mental, Neurological, and Substance Use Disorders: Key Messages from Disease Control Priorities, 3rd Edition". In: *LANCET* 387.10028 (Apr. 2016), pp. 1672–1685. ISSN: 0140-6736. DOI: 10.1016/S0140-6736(15)00390-6 (cit. on p. 155).

[Pat18a] C Patterson. "World Alzheimer Report 2018: The State of the Art of Dementia Research". In: *New frontiers. London: Alzheimer's Disease International.* (2018) (cit. on p. 90).

[Pat18b] Christina Patterson. *World Alzheimer Report 2018*. Report. Alzheimer's Disease International, Sept. 2018. (Visited on 03/27/2024) (cit. on p. 129).

[Ped+11] Fabian Pedregosa et al. "Scikit-Learn: Machine Learning in Python". In: *MACHINE LEARNING IN PYTHON* (2011) (cit. on pp. 61, 79, 95).

[Pen+22] Liling Peng et al. "Alternation of the Rich-Club Organization of Individual Brain Metabolic Networks in Parkinson's Disease". In: *Frontiers in Aging Neuroscience* 14 (2022). ISSN: 1663-4365. (Visited on 10/06/2023) (cit. on pp. 119, 127).

[Pen+11a] William D Penny et al. *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Elsevier, 2011 (cit. on p. 143).

[Pen+11b] William D. Penny et al. *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Elsevier, Apr. 2011. ISBN: 978-0-08-046650-7 (cit. on pp. 56, 77, 131).

[Pet+10] R. C. Petersen et al. "Alzheimer's Disease Neuroimaging Initiative (ADNI): Clinical Characterization". In: *Neurology* 74.3 (Jan. 2010), pp. 201–209. ISSN: 0028-3878, 1526-632X. DOI: 10.1212/WNL.0b013e3181cb3e25. (Visited on 10/13/2023) (cit. on p. 114).

[Pet+09] Ronald C. Petersen et al. "Mild Cognitive Impairment: Ten Years Later". In: *Archives of Neurology* 66.12 (Dec. 1, 2009). ISSN: 0003-9942. DOI: 10.1001/archneurol.2009.266 (cit. on p. 29).

[Phe+79] M. E. Phelps et al. "Tomographic Measurement of Local Cerebral Glucose Metabolic Rate in Humans with (F-18)2-Fluoro-2-Deoxy-D-glucose: Validation of Method". In: *Annals of Neurology* 6.5 (Nov. 1979), pp. 371–388. ISSN: 0364-5134, 1531-8249. DOI: 10.1002/ana.410060502. (Visited on 10/12/2023) (cit. on p. 106).

[Pot06] Kristin Potter. "Methods for Presenting Statistical Information: The Box Plot". In: (2006) (cit. on p. 73).

[Pri+09] Joseph L. Price et al. "Neuropathology of Nondemented Aging: Presumptive Evidence for Preclinical Alzheimer Disease". In: *Neurobiology of Aging* 30.7 (July 2009), pp. 1026–1036. ISSN: 01974580. DOI: 10.1016/j.neurobiolaging.2009.04.002 (cit. on p. 29).

[Pri+] Martin Prince et al. "World Alzheimer Report 2015. The Global Impact of Dementia: An Analysis of Prevalence, Incidence, Cost and Trends." In: () (cit. on p. 19).

[Qin+22] Kun Qin et al. "Using Graph Convolutional Network to Characterize Individuals with Major Depressive Disorder across Multiple Imaging Sites". In: *eBioMedicine* 78 (Apr. 2022), p. 103977. ISSN: 23523964. DOI: 10.1016/j.ebiom.2022.103977. (Visited on 03/14/2024) (cit. on p. 161).

[QKV09] Chengxuan Qiu et al. "Epidemiology of Alzheimer's Disease: Occurrence, Determinants, and Strategies toward Intervention". In: *Dialogues in Clinical Neuroscience* 11.2 (June 30, 2009), pp. 111–128. ISSN: 1958-5969. DOI: 10.31887/DCNS.2009.11.2/cqiu (cit. on p. 27).

[Rah+20] Farzaneh Rahmani et al. "Metabolic Connectivity in Alzheimer's Diseases". In: *Clinical and Translational Imaging* 8.3 (June 2020), pp. 157–166. ISSN: 2281-5872, 2281-7565. DOI: 10.1007/s40336-020-00371-3. (Visited on 10/12/2023) (cit. on pp. 118, 123).

[Ram+03] H. Ramaroson et al. "Prevalence of dementia and Alzheimer's disease among subjects aged 75 years or over: updated results of the PAQUID cohort". In: *Revue Neurologique* 159.4 (Apr. 2003), pp. 405–411. ISSN: 0035-3787. pmid: 12773869 (cit. on p. 27).

[Ray+22] Adwitiya Ray et al. "Artificial Intelligence and Psychiatry: An Overview". In: *Asian Journal of Psychiatry* 70 (Apr. 2022), p. 103021. ISSN: 1876-2018. DOI: 10.1016/j.ajp.2022.103021. (Visited on 03/13/2024) (cit. on p. 155).

[Rei+07] Jaap C. Reijneveld et al. "The Application of Graph Theoretical Analysis to Complex Networks in the Brain". In: *Clinical Neurophysiology* 118.11 (Nov. 2007), pp. 2317–2331. ISSN: 13882457. DOI: 10.1016/j.clinph.2007.08.010. (Visited on 10/12/2023) (cit. on p. 104).

[RBM11] Christiane Reitz et al. "Epidemiology of Alzheimer Disease". In: *Nature Reviews Neurology* 7.3 (Mar. 2011), pp. 137–152. ISSN: 1759-4758, 1759-4766. DOI: 10.1038/nrneurol.2011.2 (cit. on p. 27).

[Rém+15] Florence Rémy et al. "White Matter Disruption at the Prodromal Stage of Alzheimer's Disease: Relationships with Hippocampal Atrophy and Episodic Memory Performance". In: *NeuroImage: Clinical* 7 (2015), pp. 482–492. ISSN: 22131582. DOI: 10.1016/j.nicl.2015.01.014 (cit. on p. 28).

[Ren21] Shuhua Ren. "Metabolic Brain Network and Surgical Outcome in Temporal Lobe Epilepsy: A Graph Theoretical Study Based on 18F-fluorodeoxyglucose PET". In: (2021) (cit. on pp. 118, 123, 127).

[Rob+08] Philippe H. Robert et al. "Importance of Lack of Interest in Patients with Mild Cognitive Impairment". In: *The American Journal of Geriatric Psychiatry* 16.9 (Sept. 2008), pp. 770–776. DOI: 10.1097/jgp.0b013e31817e73db (cit. on p. 85).

[RJT15] Edmund T. Rolls et al. "Implementation of a New Parcellation of the Orbitofrontal Cortex in the Automated Anatomical Labeling Atlas". In: *NeuroImage* 122 (Nov. 2015), pp. 1–5. DOI: 10.1016/j.neuroimage.2015.07.075 (cit. on pp. 77, 131).

[RS10] Mikail Rubinov et al. "Complex Network Measures of Brain Connectivity: Uses and Interpretations". In: *NeuroImage* 52.3 (Sept. 2010), pp. 1059–1069. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2009.10.003. (Visited on 10/12/2023) (cit. on pp. 107, 109, 112, 124).

[Sad+15] Saeed Sadigh-Eteghad et al. "Amyloid-Beta: A Crucial Factor in Alzheimer's Disease". In: *Medical Principles and Practice* 24.1 (2015), pp. 1–10. ISSN: 1011-7571, 1423-0151. DOI: 10.1159/000369101 (cit. on p. 28).

[SMG94] Gopal B. Saha et al. "Radiopharmaceuticals for Brain Imaging". In: *Seminars in Nuclear Medicine*. Correlative Imaging: Part II 24.4 (Oct. 1, 1994), pp. 324–349. ISSN: 0001-2998. DOI: 10.1016/S0001-2998(05)80022-4 (cit. on p. 36).

[Sal+10] D. Salas-Gonzalez et al. "Feature Selection Using Factor Analysis for Alzheimer's Diagnosis Using PET Images". In: *Medical Physics* 37.11 (2010), pp. 6084–6095. ISSN: 2473-4209. DOI: 10.1118/1.3488894. (Visited on 03/27/2024) (cit. on p. 61).

[Sar+10] Marie Sarazin et al. "The Amnestic Syndrome of Hippocampal Type in Alzheimer's Disease: An MRI Study". In: *Journal of Alzheimer's Disease* 22.1 (Sept. 28, 2010), pp. 285–294. ISSN: 18758908, 13872877. DOI: 10.3233/JAD-2010-091150 (cit. on p. 32).

[Sch+17] Alexander Schaefer et al. "Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI". In: *Cerebral Cortex* 28.9 (July 2017), pp. 3095–3114. DOI: 10.1093/cercor/bhx179 (cit. on p. 91).

[Sch+18] Alexander Schaefer et al. "Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI". In: *Cerebral Cortex* 28.9 (Sept. 2018), pp. 3095–3114. ISSN: 1047-3211, 1460-2199. DOI: 10.1093/cercor/bhx179. (Visited on 11/10/2023) (cit. on pp. 144, 158).

[Seg12] Mohamed L. Seghier. "The Angular Gyrus". In: *The Neuroscientist* 19.1 (Apr. 2012), pp. 43–61. DOI: 10.1177/1073858412440596 (cit. on p. 85).

[Sen08a] Pavel Senin. "Dynamic Time Warping Algorithm Review". In: *Information and Computer Science Department University of Hawaii at Manoa Honolulu, USA* 855.1-23 (2008), p. 40 (cit. on p. 131).

[Sen08b] Pavel Senin. "Dynamic Time Warping Algorithm Review". In: *Information and Computer Science Department University of Hawaii at Manoa Honolulu, USA* 855.1-23 (2008), p. 40 (cit. on p. 158).

[Sha+08] David W. Shattuck et al. "Construction of a 3D Probabilistic Atlas of Human Cortical Structures". In: *NeuroImage* 39.3 (Feb. 2008), pp. 1064–1080. DOI: 10.1016/j.neuroimage.2007.09.031 (cit. on p. 90).

[She04] Z.-X Shen. "Brain Cholinesterases: III. Future Perspectives of AD Research and Clinical Practice". In: *Medical Hypotheses* 63.2 (Jan. 2004), pp. 298–307. ISSN: 03069877. DOI: 10.1016/j.mehy.2004.03.001 (cit. on p. 28).

[Shu00] Kenneth I Shulman. "Clock-Drawing: Is It the Ideal Cognitive Screening Test?" In: *International Journal of Geriatric Psychiatry* 15.6 (June 2000), pp. 548–561. ISSN: 0885-6230, 1099-1166. DOI: 10.1002/1099-1166(200006)15:6<548::AID-GPS242>3.0.CO;2-U (cit. on p. 31).

[Sil+19] Marcos Vinícius Ferreira Silva et al. "Alzheimer's Disease: Risk Factors and Potentially Protective Measures". In: *Journal of Biomedical Science* 26.1 (Dec. 2019), p. 33. ISSN: 1423-0127. DOI: 10.1186/s12929-019-0524-y (cit. on p. 29).

[Sma+15] Nadja Smailagic et al. "18F-FDG PET for the Early Diagnosis of Alzheimer's Disease Dementia and Other Dementias in People with Mild Cognitive Impairment (MCI)". In: *Cochrane Database of Systematic Reviews* (Jan. 2015). DOI: 10.1002/14651858.cd010632.pub2 (cit. on p. 85).

[SMJ14] Reisa Sperling et al. "The Evolution of Preclinical Alzheimer's Disease: Implications for Prevention Trials". In: *Neuron* 84.3 (Nov. 2014), pp. 608–622. ISSN: 08966273. DOI: 10.1016/j.neuron.2014.10.038 (cit. on p. 30).

[Spe+09] Reisa A. Sperling et al. "Amyloid Deposition Is Associated with Impaired Default Network Function in Older Persons without Dementia". In: *Neuron* 63.2 (July 2009), pp. 178–188. ISSN: 08966273. DOI: 10.1016/j.neuron.2009.07.003 (cit. on p. 33).

[Spe+11] Reisa A. Sperling et al. "Toward Defining the Preclinical Stages of Alzheimer's Disease: Recommendations from the National Institute on Aging-Alzheimer's Association Workgroups on Diagnostic Guidelines for Alzheimer's Disease". In: *Alzheimer's & Dementia* 7.3 (May 2011), pp. 280–292. ISSN: 1552-5260, 1552-5279. DOI: 10.1016/j.jalz.2011.03.003 (cit. on pp. 20, 31).

[Spo+04] O Sporns et al. "Organization, Development and Function of Complex Brain Networks". In: *Trends in Cognitive Sciences* 8.9 (Sept. 2004), pp. 418–425. ISSN: 13646613. DOI: 10.1016/j.tics.2004.07.008. (Visited on 10/12/2023) (cit. on pp. 104, 106).

[Spo13] Olaf Sporns. "Network Attributes for Segregation and Integration in the Human Brain". In: *Current Opinion in Neurobiology* 23.2 (Apr. 2013), pp. 162–171. ISSN: 09594388. DOI: 10.1016/j.conb.2012.11.015. (Visited on 10/12/2023) (cit. on pp. 106, 107).

[Spo18] Olaf Sporns. "Graph Theory Methods: Applications in Brain Networks". In: *Dialogues in Clinical Neuroscience* 20.2 (June 2018), pp. 111–121. ISSN: 1958-5969. DOI: 10.31887/DCNS.2018.20.2/osporns. (Visited on 11/06/2023) (cit. on pp. 124, 138, 155).

[STK05] Olaf Sporns et al. "The Human Connectome: A Structural Description of the Human Brain". In: *PLoS Computational Biology* 1.4 (2005), e42. ISSN: 1553-734X, 1553-7358. DOI: 10.1371/journal.pcbi.0010042. (Visited on 10/12/2023) (cit. on pp. 104, 106).

[Sta14] Cornelis J. Stam. "Modern Network Science of Neurological Disorders". In: *Nature Reviews Neuroscience* 15.10 (Oct. 2014), pp. 683–695. ISSN: 1471-003X, 1471-0048. DOI: 10.1038/nrn3801. (Visited on 10/12/2023) (cit. on p. 104).

[ST10] Kyle Strimbu et al. "What Are Biomarkers?:" in: *Current Opinion in HIV and AIDS* 5.6 (Nov. 2010), pp. 463–466. ISSN: 1746-630X. DOI: 10.1097/COH.0b013e32833ed177 (cit. on p. 31).

[Su+14] Liang Su et al. "Cerebral Metabolism in Major Depressive Disorder: A Voxel-Based Meta-Analysis of Positron Emission Tomography Studies". In: *BMC Psychiatry* 14.1 (Nov. 2014), p. 321. ISSN: 1471-244X. DOI: 10.1186/s12888-014-0321-9. (Visited on 03/13/2024) (cit. on p. 155).

[TP08] Vanessa Taler et al. "Language Performance in Alzheimer's Disease and Mild Cognitive Impairment: A Comparative Review". In: *Journal of Clinical and Experimental Neuropsychology* 30.5 (June 4, 2008), pp. 501–556. ISSN: 1380-3395, 1744-411X. DOI: 10.1080/13803390701550128 (cit. on p. 19).

[Tan+20] M. Tanveer et al. "Machine Learning Techniques for the Diagnosis of Alzheimer's Disease". In: *ACM Transactions on Multimedia Computing, Communications, and Applications* 16.1s (Jan. 2020), pp. 1–35. DOI: 10.1145/3344998 (cit. on p. 129).

[Tan12] R. E. Tanzi. "The Genetics of Alzheimer Disease". In: *Cold Spring Harbor Perspectives in Medicine* 2.10 (Oct. 1, 2012), a006296–a006296. ISSN: 2157-1422. DOI: 10.1101/cshperspect.a006296 (cit. on p. 29).

[The+17] The Alzheimer's Disease Neuroimaging Initiative et al. "Do All Roads Lead to Rome? A Comparison of Brain Networks Derived from Inter-Subject Volumetric and Metabolic Covariance and Moment-to-Moment Hemodynamic Correlations in Old Individuals". In: *Brain Structure and Function* 222.8 (Nov. 2017), pp. 3833–3845. ISSN: 1863-2653, 1863-2661. DOI: 10.1007/s00429-017-1438-7. (Visited on 10/12/2023) (cit. on pp. 104, 123).

[Thi+14] Bertrand Thirion et al. "Which fMRI Clustering Gives Good Brain Parcellations?" In: *Frontiers in Neuroscience* 8 (July 2014). DOI: 10.3389/fnins.2014.00167 (cit. on p. 91).

[Tou+14] Paule-Joanne Toussaint et al. "Characteristics of the Default Mode Functional Connectivity in Normal Ageing and Alzheimer's Disease Using Resting State fMRI with a Combined Approach of Entropy-Based and Graph Theoretical Measurements". In: *NeuroImage* 101 (Nov. 2014), pp. 778–786. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2014.08.003 (cit. on p. 33).

[Tua+21] Pham Minh Tuan et al. "AutoEncoder-based Feature Ranking for Alzheimer Disease Classification Using PET Image". In: *Machine Learning with Applications* 6 (Dec. 2021), p. 100184. DOI: 10.1016/j.mlwa.2021.100184 (cit. on pp. 75, 77–79, 84, 85, 92, 100).

[Tua+22] Pham Minh Tuan et al. "C-Atlas: A Brain Mapping Based on FDG-PET Images for Alzheimer's Disease Diagnosis". In: *2022 RIVF International Conference on Computing and Communication Technologies (RIVF)*. Ho Chi Minh City, Vietnam: IEEE, Dec. 2022, pp. 150–155. ISBN: 978-1-66546-166-5. DOI: 10.1109/RIVF55975.2022.10013857. (Visited on 10/12/2023) (cit. on p. 124).

[Tua+23] Pham Minh Tuan et al. "Dynamic Time Warping-based Distance for Constructing Individual Brain Networks from FDG-PET Images: Application to Alzheimer's Disease". In: *2023 Twelfth International Conference on Image Processing Theory, Tools and Applications (IPTA)*. Paris, France: IEEE, Oct. 2023, pp. 1–5. ISBN: 9798350325416. DOI: 10.1109/IPTA59101.2023.10320083. (Visited on 01/10/2024) (cit. on pp. 144, 157, 158).

[Tzo+02a] N. Tzourio-Mazoyer et al. "Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain". In: *NeuroImage* 15.1 (Jan. 2002), pp. 273–289. ISSN: 1053-8119. DOI: 10.1006/nimg.2001.0978. (Visited on 03/27/2024) (cit. on p. 56).

[Tzo+02b] N. Tzourio-Mazoyer et al. "Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain". In: *NeuroImage* 15.1 (Jan. 2002), pp. 273–289. DOI: 10.1006/nimg.2001.0978 (cit. on p. 90).

[Tzo+02c] N. Tzourio-Mazoyer et al. "Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain". In: *NeuroImage* 15.1 (Jan. 2002), pp. 273–289. ISSN: 10538119. DOI: 10.1006/nimg.2001.0978. (Visited on 10/12/2023) (cit. on p. 107).

[VS19] Martijn P. Van Den Heuvel et al. "A Cross-Disorder Connectome Landscape of Brain Dysconnectivity". In: *Nature Reviews Neuroscience* 20.7 (July 2019), pp. 435–446. ISSN: 1471-003X, 1471-0048. DOI: 10.1038/s41583-019-0177-6. (Visited on 10/12/2023) (cit. on p. 123).

[van+14] Stéfan van der Walt et al. "Scikit-Image: Image Processing in Python". In: *PeerJ* 2 (June 2014), e453. ISSN: 2167-8359. DOI: 10.7717/peerj.453. (Visited on 06/14/2022) (cit. on p. 95).

[Van+10] Rik Vandenberghe et al. "^{18}F-flutemetamol Amyloid Imaging in Alzheimer Disease and Mild Cognitive Impairment: A Phase 2 Trial". In: *Annals of Neurology* 68.3 (Sept. 2010), pp. 319–329. ISSN: 0364-5134, 1531-8249. DOI: 10.1002/ana.22068 (cit. on p. 35).

[Van+13] Rik Vandenberghe et al. "Binary Classification of 18F-flutemetamol PET Using Machine Learning: Comparison with Visual Reads and Structural MRI". In: *NeuroImage* 64 (Jan. 2013), pp. 517–525. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2012.09.015 (cit. on p. 36).

[Var+12] Erdem Varol et al. "Feature Ranking Based Nested Support Vector Machine Ensemble for Medical Image Classification". In: *2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)*. May 2012, pp. 146–149. DOI: 10.1109/ISBI.2012.6235505. (Visited on 03/27/2024) (cit. on p. 52).

[VS+11] Prashanthi Vemuri et al. "Antemortem Differential Diagnosis of Dementia Pathology Using Structural MRI: Differential-STAND". In: *NeuroImage* 55.2 (Mar. 2011), pp. 522–531. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2010.12.073 (cit. on p. 32).

[Vil+18] Victor L. Villemagne et al. "Imaging Tau and Amyloid-β Proteinopathies in Alzheimer Disease and Other Conditions". In: *Nature Reviews Neurology* 14.4 (Apr. 2018), pp. 225–236. ISSN: 1759-4758, 1759-4766. DOI: 10.1038/nrneurol.2018.9 (cit. on p. 36).

[Wan10] Jinhui Wang. "Graph-Based Network Analysis of Resting-State Functional MRI". In: *Frontiers in Systems Neuroscience* (2010). ISSN: 16625137. DOI: 10.3389/fnsys.2010.00016. (Visited on 10/12/2023) (cit. on p. 104).

[Wan+18] Kai-Liang Wang et al. "Metabolic Covariance Networks Combining Graph Theory Measuring Aberrant Topological Patterns in Mesial Temporal Lobe Epilepsy". In: *CNS Neuroscience & Therapeutics* 25.3 (Oct. 2018), pp. 396–408. ISSN: 1755-5930. DOI: 10.1111/cns.13073. (Visited on 10/05/2023) (cit. on pp. 118, 123, 127).

[Wan05] Lipo Wang. *Support vector machines: Theory and applications*. Vol. 177. Springer Science & Business Media, 2005 (cit. on p. 41).

[Wan+20a] Min Wang et al. "Individual Brain Metabolic Connectome Indicator Based on Kullback-Leibler Divergence Similarity Estimation Predicts Progression from Mild Cognitive Impairment to Alzheimer's Dementia". In: *European Journal of Nuclear Medicine and Molecular Imaging* 47.12 (2020), pp. 2753–2764. ISSN: 1619-7070. DOI: 10.1007/s00259-020-04814-x. (Visited on 10/06/2023) (cit. on pp. 113, 120, 121, 123, 124, 127).

[Wan+20b] Min Wang et al. "Individual Brain Metabolic Connectome Indicator Based on Kullback-Leibler Divergence Similarity Estimation Predicts Progression from Mild Cognitive Impairment to Alzheimer's Dementia". In: *European Journal of Nuclear Medicine and Molecular Imaging* 47.12 (Apr. 2020), pp. 2753–2764. DOI: 10.1007/s00259-020-04814-x (cit. on pp. 129, 132).

[Wan+20c] Min Wang et al. "Individual Brain Metabolic Connectome Indicator Based on Kullback-Leibler Divergence Similarity Estimation Predicts Progression from Mild Cognitive Impairment to Alzheimer's Dementia". In: *European Journal of Nuclear Medicine and Molecular Imaging* 47.12 (Apr. 2020), pp. 2753–2764. DOI: 10.1007/s00259-020-04814-x (cit. on pp. 144, 157).

[Wan+23] Xiaoyang Wang et al. "Abnormal Metabolic Connectivity in Default Mode Network of Right Temporal Lobe Epilepsy". In: *Frontiers in Neuroscience* 17 (2023). ISSN: 1662-453X. (Visited on 10/05/2023) (cit. on pp. 118, 119, 127).

[WS98] Duncan J. Watts et al. "Collective Dynamics of 'Small-World' Networks". In: *Nature* 393.6684 (June 1998), pp. 440–442. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/30918. (Visited on 10/12/2023) (cit. on p. 107).

[Węg18a] Stanisław Węglarczyk. "Kernel Density Estimation and Its Application". In: *ITM Web of Conferences* 23 (2018). Ed. by W. Zielinski et al., p. 00037. DOI: 10.1051/itmconf/20182300037 (cit. on p. 131).

[Węg18b] Stanisław Węglarczyk. "Kernel Density Estimation and Its Application". In: *ITM Web of Conferences* 23 (2018). Ed. by W. Zielinski et al., p. 00037. DOI: 10.1051/itmconf/20182300037 (cit. on pp. 144, 158).

[WQG85] C W Wong et al. "Neuritic Plaques and Cerebrovascular Amyloid in Alzheimer Disease Are Antigenically Related." In: *Proceedings of the National Academy of Sciences* 82.24 (Dec. 1985), pp. 8729–8732. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.82.24.8729 (cit. on p. 28).

[Xu+19] Keyulu Xu et al. *How Powerful Are Graph Neural Networks?* Feb. 2019. arXiv: 1810.00826 [cs, stat]. (Visited on 03/13/2024) (cit. on pp. 158, 159).

[YDH17] Igor Yakushev et al. "Metabolic Connectivity: Methods and Applications". In: *Current Opinion in Neurology* 30.6 (Dec. 2017), pp. 677–685. ISSN: 1350-7540, 1473-6551. DOI: 10.1097/WCO.0000000000000494. (Visited on 10/12/2023) (cit. on p. 118).

[Yao+15] Zhijun Yao et al. "A FDG-PET Study of Metabolic Networks in Apolipoprotein E *E*4 Allele Carriers". In: *PLOS ONE* 10.7 (July 2015), e0132300. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0132300. (Visited on 10/06/2023) (cit. on pp. 120, 121, 127).

[Yao+16] Zhijun Yao et al. "Individual Metabolic Network for the Accurate Detection of Alzheimer's Disease Based on FDGPET Imaging". In: *2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. Dec. 2016, pp. 1328–1335. DOI: 10.1109/BIBM.2016.7822710. (Visited on 10/06/2023) (cit. on pp. 113, 120, 121, 124, 127).

[Yao+17] Zhijun Yao et al. "Learning Metabolic Brain Networks in MCI and AD by Robustness and Leave-One-out Analysis: An FDG-PET Study". In: *American Journal of Alzheimer's Disease & Other Dementias* 33.1 (Sept. 2017), pp. 42–54. DOI: 10.1177/1533317517731535 (cit. on p. 129).

[Yao+18] Zhijun Yao et al. "Learning Metabolic Brain Networks in MCI and AD by Robustness and Leave-One-Out Analysis: An FDG-PET Study". In: *American Journal of Alzheimer's Disease & Other Dementias®* 33.1 (Feb. 2018), pp. 42–54. ISSN: 1533-3175. DOI: 10.1177/1533317517731535. (Visited on 10/06/2023) (cit. on pp. 120, 121, 127).

[Yu+23] Jie Yu et al. "Construction of an Individualized Brain Metabolic Network in Patients with Advanced Non-Small Cell Lung Cancer by the Kullback-Leibler Divergence-Based Similarity Method: A Study Based on 18F-fluorodeoxyglucose Positron Emission Tomography". In: *Frontiers in Oncology* 13 (2023). ISSN: 2234-943X. (Visited on 10/06/2023) (cit. on pp. 121, 122, 124, 127).

[Yu+11] Shan Yu et al. "Higher-Order Interactions Characterized in Cortical Activity". In: *Journal of Neuroscience* 31.48 (Nov. 2011), pp. 17514–17526. ISSN: 0270-6474, 1529-2401. DOI: 10.1523/JNEUROSCI.3127-11.2011. (Visited on 03/13/2024) (cit. on p. 156).

[Zha+20] Li Zhang et al. "A Survey on Deep Learning for Neuroimaging-Based Brain Disorder Analysis". In: *Frontiers in Neuroscience* 14 (Oct. 2020), p. 779. ISSN: 1662-453X. DOI: 10.3389/fnins.2020.00779. (Visited on 03/13/2024) (cit. on pp. 138, 139, 155).

[Zha+07] Y. Zhang et al. "Diffusion Tensor Imaging of Cingulum Fibers in Mild Cognitive Impairment and Alzheimer Disease". In: *Neurology* 68.1 (Jan. 2, 2007), pp. 13–19. ISSN: 0028-3878, 1526-632X. DOI: 10.1212/01.wnl.0000250326.77323.01 (cit. on p. 34).

[Zha+17] Yu Zhao et al. "Predicting MCI Progression with Individual Metabolic Network Based on Longitudinal FDG-PET". In: *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. Nov. 2017, pp. 1894–1899. DOI: 10.1109/BIBM.2017.8217948. (Visited on 10/06/2023) (cit. on pp. 120, 121, 127).

[ZA10] Karl Zilles et al. "Centenary of Brodmann's Map—Conception and Fate". In: *Nature Reviews Neuroscience* 11.2 (2010), pp. 139–145 (cit. on p. 91).

[Zvě19] Martina Zvěřová. "Clinical Aspects of Alzheimer's Disease". In: *Clinical Biochemistry*. Alzheimer's Disease: Making the Point 72 (Oct. 1, 2019), pp. 3–6. ISSN: 0009-9120. DOI: 10.1016/j.clinbiochem.2019.04.015 (cit. on p. 19).