Neuroanatomical-Based Machine Learning Prediction of Alzheimer's Disease Across Sex and Age

by

Bhaavin Kishore Jogeshwar

A Thesis

Submitted to the Faculty

of the

WORCESTER POLYTECHNIC INSTITUTE

In partial fulfillment of the requirements for the

Degree of Master of Science

in

Robotics Engineering

July 2024

APPROVED:

Professor Benjamin Nephew, Thesis Co-Advisor, Biology & Biotechnology Dept., WPI

Professor Loris Fichera, Thesis Co-Advisor, Robotics Engineering Dept., WPI

Professor Haichong Zhang, Thesis Committee Member, Robotics Engineering, WPI

Abstract

Alzheimer's Disease (AD) is a progressive neurodegenerative disorder characterized by cognitive decline and memory loss. In 2024, in the US alone, it affects approximately 1 in 9 people aged 65 and older, representing 10.9% of this population. This amounts to 6.9 million individuals, with women (4.2 million) constituting more than men (2.7 million). Magnetic resonance imaging (MRI) has emerged as a valuable tool for examining brain structure and identifying potential AD biomarkers. Early detection and accurate AD diagnosis are crucial for timely intervention and management. Moreover, monitoring disease progression and evaluating treatment effectiveness heavily rely on identifying reliable biomarkers. While sex contributes to Alzheimer's prevalence, age remains the primary risk factor, with incidence increasing significantly with each decade. However, the reasons for the variation of biomarkers with age remain unclear. This study performs predictive analyses by employing machine learning techniques to identify key brain regions associated with AD using numerical data derived from anatomical MRI scans, going beyond standard statistical methods. Additionally, subgroup analyses identified key brain regions that strongly predicted AD across three age groups: younger (69-76 years), older (77-84 years), and unified (69-84 years). Using the Random Forest Algorithm, we achieved 92.87% accuracy in detecting AD from Mild Cognitive Impairment, and Cognitive Normals. The hippocampus, amygdala, and entorhinal cortex consistently showed volume decreases across sexes and age groups despite varying prevalence rates between males and females. For instance, the right amygdala exhibited decreased volume in younger males (aged 69-76), while in females, this decline was observed in the older group (aged 77-84). Both younger males and females (aged 69-76) exhibited

volume decreases in the right hippocampus, suggesting its importance in the early stages of AD. Older males (aged 77-84) showed substantial volume decreases in the left inferior temporal cortex. Additionally, the left middle temporal cortex showed decreased volume in females, suggesting a potential female-specific influence, while the right entorhinal cortex may have a male-specific impact. These age-specific sex differences could inform clinical research and treatment strategies, aiding in identifying neuroanatomical markers and therapeutic targets for future research and clinical interventions.

Keywords: Machine Learning, Random Forest, Alzheimer's Disease, anatomical MRI, ADNI, neuroimaging, age group, sex, FastSurfer, FreeSurfer, Clinica, trends.

Acknowledgements

I am deeply grateful to Professor Benjamin Nephew for providing me with the opportunity to work with him on this project. Working under your guidance has significantly enhanced my skills and shaped me into a better scientist.

I extend my thanks to Professor Loris Fichera for facilitating my collaboration in neuroscience with Professor Nephew and for guiding me in academic matters. Along with him, I appreciate Professor Haichong Zhang for supporting the committee formation. I am thankful to Dr. Marcelo Febo for his invaluable insights and to Senbao Lu for his constructive feedback on Machine Learning concepts.

I profoundly appreciate Dr. Anjali Jogeshwar, and Dr. Nuzhet Nihaar Hafiza Nasir Ahamed, and Meenal Nandwani-Gawdi for always being there whenever I needed them and for celebrating with me every success, big or small, throughout this project. My heartfelt thanks go to The Global Lab, WPI Squash Club, and my friends—Mayank Bansal, Vishrut Bohara, Miheer Diwan, Aabha Tamhankar, Girivaasan Chandrasekaran, Sanya Gulati, Shreya Bang, and Aayushee Patel—for their support, and unwavering faith in me.

Lastly, I am grateful for the immense wellspring of love and encouragement from my family in India, making this journey (12000 km away from home) possible.

This thesis is dedicated to my parents, Ritu and Kishore Jogeshwar, and my sister, Dr. Anjali Jogeshwar, for being my pillars of strength and for their endless love and support.

Contents

1 Introduction **1**

1.1 Prevalence of AD 1

1.2 Early Challenges of AD 2

1.3 fMRI and sMRI 3

1.4 ADNI Dataset 5

1.5 Supervised Learning in Neuro 6

1.6 Subgroup Analyses in AD 8

1.7 Clinical Trials 9

1.7.1 Past Clinical Trials of Alzheimer's Disease 9

1.7.2 Reasons for Trial Failures 10

1.7.3 Recent Studies, Trials, and Approved Drugs 12

1.8 Future of AD 13

2 Methodology **17**

2.1 ADNI Dataset 17

2.2 Preprocessing 19

2.2.1 Clinica 19

2.2.2 FreeSurfer 22

2.2.3 FastSurfer 23

2.3 Class Imbalance 25
2.4 Random Forest 26
2.5 RF Neuroanatomical Analysis 30
2.5.1 Understanding Gini Impurity 30
2.5.2 Feature Importance 30
2.6 Model Validation 31
2.6.1 K-Fold Cross-Validation 31
2.6.2 Stratified K-Fold Cross-Validation 32
2.6.3 Leave-One-Out Cross-Validation 32
2.6.4 Implementation 33
2.7 Performance Metrics 33

3 Results **36**
3.1 Clinica outputs 36
3.2 FastSurfer outputs 37
3.3 Feature Engineering 39
3.4 Subgroup Analyses 40
3.4.1 Statistical Analysis of Brain Regions 40
3.4.2 Performance Metrics 44
3.4.3 RF Neuroanatomical Analysis 44
3.4.4 Neuroanatomical Trends in AD 48
3.4.5 Insights from Visual Representations 51

4 Discussion **56**
4.1 Markers of AD 57
4.1.1 Role of hippocampus and amygdala 57
4.1.2 Role of entorhinal cortex 58

4.2 Sex and Age Subgroup Analyses 60
4.3 Medications for AD 61
4.3.1 Symptomatic Medications 62
4.3.2 Disease-Modifying Drugs (Under Investigation) 63
4.4 Challenges Faced 66
4.4.1 Dataset to BIDS Hierarchy Transformation 66
4.4.2 Building the Machine Learning Dataset 67
4.5 Future Work 68

5 Supplementary Material **72**

Bibliography **76**

List of Figures

2.1 Conversion to BIDS Architecture. 22

2.2 An example flowchart of Random Forest classifier 27

3.1 Defaced sMRI scans 37

3.2 Brain parcellation in Freeview 38

3.3 Average volumes of top brain regions for statistical analyses 41

3.4 Average volumes of four brain regions 42

3.5 All occurrences of brain regions 52

3.6 Most common high-ranking brain regions 53

3.7 Visualization of unique brain regions using Freeview 54

3.8 Top 6 occurrences of brain regions 55

3.9 Visualization of top 6 brain regions on Freeview 55

4.1 Comparison of a discarded scan with a high-quality scan 68

4.2 Segmentation of brain tissues using FSLeyes 70

List of Tables

2.1 ADNI demographic characteristics 18

2.2 Distribution of ADNI Participants by Age, Sex, and Diagnosis 19

2.3 Comparison of FreeSurfer and FastSurfer 24

2.4 Comparison of FreeSurfer and FastSurfer Analyses on Different Machines 25

2.5 Processing Time for Volumetric and Thickness Analysis 25

3.1 P-values before Benjamini-Hochberg correction 43

3.2 P-values after Benjamini-Hochberg correction 43

3.3 Comparison of Aggregated Performance Metrics using K-fold Cross-Validation 44

3.4 Comparison of Aggregated Performance Metrics using Stratified K-fold Cross-Validation 45

3.5 Comparison of Aggregated Performance Metrics using Leave-One-Out Cross-Validation 45

3.6 Comparison of Top Contributing Features using K-fold Cross-Validation 46

3.7 Comparison of Top Contributing Features using Stratified K-fold Cross-Validation 46

3.8 Comparison of Top Contributing Features using Leave-One-Out Cross-Validation 47
3.9 Consistent Top Contributing Features using K-fold, Stratified K-fold, and Leave-One-Out Cross-Validation 48

5.1 Comparison of Performance Metrics and Top Contributing Features using K-fold Cross-Validation 74
5.2 Comparison of Performance Metrics and Top Contributing Features using Stratified K-fold Cross-Validation 74
5.3 Comparison of Performance Metrics and Top Contributing Features using Leave-One-Out Cross-Validation 75

Chapter 1

Introduction

Alzheimer's Disease (AD) is a type of dementia that affects millions of individuals worldwide, particularly the elderly population. It leads to a gradual loss of cognitive abilities, and as the condition advances, individuals require more assistance with daily tasks. Currently, AD remains the fifth-leading cause of death among Americans aged 65 and older [1]. There is an urgent need for continued research and intervention strategies to address the growing impact of AD on individuals, families, and society.

1.1 Prevalence of AD

The prevalence of AD refers to the proportion of individuals within a specific population affected by the condition at a given time. Recent estimates indicate a global increase in the incidence and prevalence of AD, with a 147.95% rise in incidence and a 160.84% increase in prevalence observed from 1990 to 2019 [2]. Crude AD prevalence during 2015–2020 was estimated at 760.5 per 100,000 inhabitants, indicating a substantial burden [3]. The age-standardized rates for incidence, prevalence, death, and disability-adjusted life years consistently rose during this period, impacting both

men and women [2]. This upward trend is further emphasized by the findings of a study focused on South Korea, reporting an overall increase in AD incidence and prevalence from 2006 to 2015, particularly in the population aged 40 years or older [4].

According to 2023 estimates, approximately 6.7 million Americans aged 65 and older are living with AD [1]. This number is projected to reach 13.8 million by 2060 [5] unless remarkable medical breakthroughs to prevent, slow, or cure the disease are made. Among MCI subjects, only 10–15% develop AD annually while others will remain cognitively stable [6]. Additionally, a comprehensive investigation spanning from the pre-onset phase to advanced stages revealed that up to 90% of dementia subjects, including AD, experience Neuropsychiatric or Behavioral and Psychological Symptoms of Dementia (BPSD) [7].

The economic burden of AD is substantial, with a meta-analysis revealing an estimated total cost of $20,461 per patient per year. The breakdown according to disease stages indicates costs of $14,675 for the mild stage, $19,975 for the moderate stage, and $29,708 for the severe stage [8]. These findings collectively underscore the growing global burden of AD, emphasizing the need for targeted interventions and policy initiatives to address its escalating incidence and prevalence.

1.2 Early Challenges of AD

The early challenges of AD are primarily related to cognitive function. Brain changes related to AD may begin at least 20 years before symptoms appear [9]. Individuals experience mild memory lapses in the initial stages, such as forgetting names, appointments, or where they placed everyday items. These early cognitive difficulties may be dismissed as normal aging or stress.

As the disease progresses, challenges intensify, leading to more noticeable symptoms. During the intermediate stages of AD, cognitive decline becomes more apparent. Individuals may struggle with everyday tasks, such as following a conversation, organizing activities, or even managing finances. Memory loss becomes more pronounced and may extend to forgetting important events, faces, or personal history. Behavioral changes, mood swings, and difficulty with problem-solving can also emerge.

In the advanced stages of AD, individuals face severe cognitive impairment and a profound loss of functional abilities. Memory loss becomes profound, and individuals may not recognize close family members. Communication skills decline predominantly, and individuals may struggle to articulate thoughts or comprehend language. Basic self-care tasks become increasingly challenging, and individuals often require round-the-clock care.

Throughout these stages, the progression of AD is marked by the accumulation of abnormal protein deposits in the brain, such as beta-amyloid plaques and tau tangles [10], [11]. These deposits disrupt communication between brain cells and lead to the widespread death of neurons, contributing to the cognitive decline [12], [13] observed in individuals with AD. The gradual worsening of symptoms is an indication of the degenerative nature of the disease and the need for ongoing research to understand its underlying mechanisms and develop effective interventions.

1.3 fMRI and sMRI

Functional Magnetic Resonance Imaging (fMRI) and Structural Magnetic Resonance Imaging (sMRI) are two advanced neuroimaging techniques that play crucial roles in understanding the intricacies of the human brain. fMRI focuses on capturing

changes in blood flow and helps identify brain regions activated during specific tasks or stimuli. In this method, neural activity is inferred based on blood oxygenation level-dependent (BOLD) signals, providing insights into the functional aspects of the brain. sMRI provides detailed structural information about the brain's anatomy. sMRI focuses on the brain's structure, including gray and white matter* and cerebrospinal fluid, allowing researchers to analyze abnormalities, measure volumes, and study anatomical connections.

In our research, we have chosen to focus on sMRI due to its less variable data and the less complex and subjective analysis compared to fMRI. For example, sMRI captures structural information that, while subject to gradual changes over time, remains relatively stable over shorter intervals, while fMRI measures brain activity that can vary with the participant's immediate thoughts or tasks. This stability with sMRI allows for more reliable and consistent results in studies of neurodegenerative conditions. Additionally, the analysis of sMRI is more straightforward, focusing on measurable structural changes in the brain, such as volume and thickness, rather than the complex interpretation of brain activity patterns seen in fMRI. This makes sMRI a robust tool for examining the consistent structural changes associated with neurological disorders. The targeted approach in sMRI, concentrating on smaller yet prominent brain regions, provides a nuanced perspective that can contribute to advancements in understanding various neurodegenerative conditions, including Alzheimer's Disease, Parkinson's Disease, Schizophrenia, and bipolar disorder.

*Gray matter contains neuronal cell bodies and dendrites, which enable communication among neighboring neurons and serve as the central hub for processing information related to sensation, perception, movement, learning, speech, and cognition [14]. Conversely, white matter consists of axons that transmit impulses between distant brain regions and the spinal cord. Its primary function is to facilitate communication between different areas of gray matter and connect gray matter with the rest of the body [15].

1.4 ADNI Dataset

The Alzheimer's Disease Neuroimaging Initiative (ADNI) [16] is a longitudinal multicenter study aimed at developing clinical, imaging, genetic, and biochemical biomarkers for the early detection and tracking of AD. The first phase of ADNI (ADNI-1) commenced in October 2004 with a duration of 5 years and funding totaling $67 million, including contributions from the National Institute on Aging, pharmaceutical companies, and foundations. Its primary goal was to develop biomarkers as outcome measures for clinical trials. ADNI-1's study included 200 elderly cognitive normals (CN), 400 individuals with mild cognitive impairment (MCI), and 200 subjects with early AD. Through the analysis of brain scans, genetic profiles, and biomarkers from blood and cerebrospinal fluid, ADNI-1 aimed to identify more sensitive and accurate biomarkers for the early detection and tracking of AD progression. As part of the study, brain imaging measures such as structural MRI and PET scans were utilized to assess glucose metabolism and amyloid accumulation in the brain. This initiative supports advances in AD intervention, prevention, and treatment by facilitating early diagnosis and tracking of disease progression with an innovative data-access policy enabling worldwide access to its findings.

To ensure the credibility and reliability of this study, we utilized the ADNI database. This database comprises MRI scans from individuals diagnosed with AD, MCI, and CN conditions. We sourced raw and preprocessed samples directly from the ADNI website for our neuroimaging software to maintain quality and comparability across research endeavors. ADNI is a valuable resource that aligns with our study's objectives of investigating biomarkers associated with AD progression.

1.5 Supervised Learning in Neuroimaging

Machine learning (ML), a subfield of artificial intelligence (AI), plays a crucial role in neuroimaging and medical research by developing algorithms capable of learning from data to make predictions or decisions. ML algorithms are broadly categorized into supervised and unsupervised learning approaches.

Supervised learning involves training algorithms on labeled datasets, where each data point is associated with a known outcome. These labeled datasets guide algorithms to learn patterns and relationships, enabling them to make predictions or classifications when presented with new, unseen data. In contrast, unsupervised learning algorithms analyze unlabeled data to identify inherent patterns, structures, or relationships without explicit guidance from labeled examples.

In neuroimaging, supervised learning is essential for analyzing MRI scans and identifying biomarkers associated with AD and other neurological conditions. By using labeled datasets where each MRI scan is annotated with its corresponding diagnosis, supervised learning algorithms can learn to recognize patterns indicative of disease progression or pathology.

Supervised ML offers several advantages over traditional statistical methods in neuroimaging and medical research. For example, when studying MRI scans for AD neuroanatomical markers, supervised ML algorithms can efficiently learn from labeled data, such as brain structure changes linked to disease progression. Meanwhile, traditional statistical methods, such as linear regression or analysis of variance (ANOVA), may struggle to capture nonlinear relationships or complex interactions between brain regions. Supervised ML techniques, such as support vector machines or random forests, allow researchers to use diverse imaging features, such as voxel intensities, to enhance diagnosis accuracy and better understand neuro-

logical conditions, improving patient care. Moreover, supervised learning enables researchers and clinicians to develop predictive models for early disease detection, treatment response prediction, and patient outcome prognosis. These models use features extracted from MRI scans to predict disease onset, progression, or treatment effectiveness [17].

In recent years, advancements in neuroimaging analysis have underscored the potential of these techniques in early AD detection and clinical progression prediction [18], [19]. The review by Ahmadzadeh et al. (2023) [20] consolidates data from various modalities, including positron emission tomography (PET), structural magnetic resonance imaging (sMRI), functional magnetic resonance imaging (fMRI), single photon emission computed tomography (SPECT), and electroencephalography (EEG), commonly used for acquiring both functional and anatomical brain data. Some researchers have explored combinations of these modalities to assemble a more comprehensive neurological dataset, addressing both advantages and limitations. However, the integration of multiple measures and modalities associated with the transition to AD dementia presents ongoing challenges due to the brain's complex nature.

The current importance of machine learning in clinical practice lies in its ability to detect, measure, and compare disease-related patterns. There are three key directions in which the field is progressing [17]. Firstly, machine learning methods can help link observed phenotypic data from neuroimaging to underlying biological mechanisms, bridging the gap between imaging data and molecular markers. Secondly, embedding methods allow researchers to study brain architecture more thoroughly, especially during disease progression and treatment response. This helps explore the complex relationship between anatomy and function. Lastly, using finer-scale analyses with machine learning can provide deeper insights into small varia-

tions in brain structure, making distinguishing disease-related changes from normal ones more accurate. These advancements could lead to earlier detection and more targeted interventions for neurological conditions.

1.6 Subgroup Analyses in AD

Subgroup analyses are conducted to understand the heterogeneity of AD and identify potential neuroanatomical markers associated with the disease. In this project, we divide the dataset into sex-specific and age-specific groups to uncover variations in brain structure across different demographic segments. These analyses aim to identify patterns that may predict disease presence and contribute to a more personalized understanding of AD.

Rationale for Sex and Age-Specific Predictions

1. **Sex Differences**: Females tend to have a higher prevalence of AD compared to males, potentially due to hormonal differences and genetic factors [21], [22]. The lifetime risk for developing AD at age 45 is 1 in 5 for women and 1 in 10 for men [23]. As of 2024, out of the 6.9 million individuals aged 65 and older with AD, 4.2 million are women, and 2.7 million are men [23], [24]. It is unknown whether males and females may benefit differently from age, sex, and neuroanatomically specific treatments. These statistics highlight the importance of exploring sex-specific predictions in AD research. Exploring sex-specific predictions allows us to investigate potential biological, genetic, and hormonal factors contributing to these differences.

2. **Age Variability**: Age is a well-established risk factor for AD, with the incidence of the disease increasing with advancing age. The "age of onset" refers to the age at which symptoms of a condition first appear in an individual.

The age of onset for AD is generally around 65 years [25]. In the US, approximately 110 out of every 100,000 adults aged 30 to 64 have young-onset Alzheimer's, defined as AD occurring before the age of 65. [25]. Analyzing age-specific predictions can help determine if MCI and AD neuroanatomical predictors vary with age. Understanding these variations can help identify specific age groups that may be particularly vulnerable to AD onset and aid in developing age-specific interventions.

Subgroup analyses reveal detailed patterns in the data, improving our understanding of AD for tailored diagnosis and treatment.

1.7 Clinical Trials

1.7.1 Past Clinical Trials of Alzheimer's Disease

AD has been the focus of extensive research efforts to understand its underlying mechanisms and develop effective treatments. Over the decades, various hypotheses have been proposed to explain the pathogenesis of AD, each guiding the design of numerous clinical trials [26]. These hypotheses shaped our understanding of the disease and influenced the development of therapeutic strategies. This section reviews essential hypotheses that have driven past clinical trials, highlighting their theoretical foundations and outcomes.

Cholinergic Hypothesis

The cholinergic hypothesis, proposed in 1976 [27], posits that AD results from reduced acetylcholine synthesis, leading to cognitive decline [28]–[30]. While extensive trials with acetylcholinesterase inhibitors showed some initial success in improving

symptoms [31]–[33], they did not halt the disease's course. Consequently, regulatory bodies downgraded the medical benefit rating of these treatments [34].

Amyloid Hypothesis

In 1991, the amyloid hypothesis identified amyloid-β deposition as a primary trigger in AD pathology [35], [36]. This hypothesis led to numerous clinical trials focusing on Aβ-targeting therapies, such as antibody vaccines and secretase inhibitors [37]. Despite promising preclinical results, these trials largely failed to translate preclinical successes into effective patient treatments [37], [38].

Tau Propagation Hypothesis

The tau propagation hypothesis, emphasized since 2009 [39], highlights the role of tau protein in forming neurofibrillary tangles, a hallmark of AD [40], [41]. Clinical trials investigating tau-targeting therapies, such as tau assembly inhibitors, have produced mixed results. While some agents showed early promise [42], [43], their overall efficacy in halting disease progression remains uncertain.

Along with others namely the mitochondrial cascade [44], calcium homeostasis [45], and neurovascular [46] hypotheses, these hypotheses continue to shape research and clinical trials in pursuing effective AD treatments. Despite decades of exploration and advancements in understanding AD's mechanisms, translating these insights into successful therapies that alter disease progression remains a formidable challenge.

1.7.2 Reasons for Trial Failures

Numerous factors have contributed to the failure of clinical trials for AD treatments. These factors can be divided into two main categories: insufficient evidence for

initiating pivotal trials and issues related to pivotal trial design [47].

Insufficient Evidence for Initiating Pivotal Trials

Phase III trials for AD often began without a solid foundation of evidence from earlier phases. This was due to:

- **Insufficient testing for clinical efficacy:** Some drugs advanced to phase III with minimal efficacy testing, often based on limited clinical studies or epidemiological observations.
- **Over-reliance on biomarker data:** Changes in biomarkers (amyloid-β or tau) were sometimes accepted as surrogate endpoints without confirmatory cognitive or functional performance changes.
- **Incorrect choice of drug dose:** Proper dosages were not permanently established in phase II trials, leading to inadequate dosing in phase III.
- **Inappropriate reliance on post hoc subgroup analyses:** Decisions to advance to phase III were sometimes based on subgroup analyses of failed trials, which produced spurious results due to smaller sample sizes and multiple statistical comparisons [48].

Pivotal Trial Design Issues

Phase III clinical trial designs often did not follow optimal drug development practices, including:

- **Poor choice of primary clinical outcome measures:** Common measures such as ADAS-Cog, developed in the 1980s, were not sensitive enough for earlier-stage AD patients. Alternative measures such as ADCOMS [49] have shown more sensitivity.

- **Inclusion of non-AD patients:** Misdiagnosis of AD, including other dementias or conditions, led to trial participants who did not have AD, complicating the demonstration of treatment efficacy.

- **Insufficient accounting for potential AD subtypes:** AD is understood to be heterogeneous, potentially encompassing several subtypes. Trials often did not account for this variability, which could affect cognitive test results and drug effectiveness.

- **Therapeutic interventions administered too late:** Amyloid-directed treatments often failed because they were started after substantial plaque buildup. Early intervention is crucial to halt neurodegenerative processes. Targeting younger or early-stage patients might be more effective despite an increased risk of misdiagnosis due to subtler symptoms.

1.7.3 Recent Studies, Trials, and Approved Drugs

Lecanemab (Leqembi) is a humanized IgG1 antibody targeting soluble amyloidβ protofibrils. It received accelerated approval from the US FDA on January 6, 2023, based on phase 2 trial evidence showing amyloid removal and potential clinical benefits [50], [51]. A double-blind, placebo-controlled phase 2 trial with 856 patients demonstrated substantial, dose-dependent amyloid plaque reduction with Lecanemab treatment (10 mg/kg biweekly) over 79 weeks [52]. Currently, three phase 3 trials are underway. These include Clarity AD (NCT03887455[‡]), which reported favorable outcomes across primary and secondary measures, AHEAD 3–45 (NCT04468659[‡]), and the DIAN-TU Next Generation trial (NCT05269394[‡]). On July 6, 2023, Lecanemab received traditional FDA approval based on phase 3 Clarity

[‡]Trial listed in clinicaltrials.gov

AD trial data [53].

Donanemab, developed from mouse mE8-IgG2a, targets amyloidβ (3–42) in amyloid plaques [54]. In the phase 2 TRAILBLAZER-ALZ study, Donanemab met its primary endpoint by delaying cognitive decline by 32% compared to placebo, correlating amyloid reduction with improved iADRS scores in *ApoE4* carriers [55]. Currently, five phase 3 trials are ongoing. TRAILBLAZER-ALZ 2 (NCT04437511‡) and TRAILBLAZER-ALZ 3 (NCT05026866‡) focus on early symptomatic AD and prevention. Results indicate substantial cognitive improvement in low/medium tau groups and a 60% decline reduction in early-stage AD, with common adverse effects including amyloid-related imaging abnormalities and infusion reactions [55], [56].

Recent AD drug trials have focused on amyloid and tau mechanisms. Despite early anti-amyloid trial failures, the positive outcomes from Lecanemab and Donanemab trials [57], [58] have reignited interest in amyloid-related therapies. Research now targets prodromal or preclinical stages to address unmet needs such as neuroprotection and anti-neuroinflammation. Future trials may increasingly incorporate amyloid-related therapies within AD treatment approaches [59], [60].

1.8 Future of AD

As we look ahead, the future of AD raises critical questions about how we can proactively address and prepare for this challenging condition. Research in this field is essential as we aim to understand how lifestyle choices, early interventions, and preventive measures influence the course of AD. There's a growing consensus that maintaining a healthy diet, staying physically active, and engaging in cognitive exercises may reduce the risk of developing AD. Early detection and intervention are essential, and ongoing research aims to find potential biomarkers and innovative

strategies for timely diagnosis and treatment. This research is crucial for individuals, families, communities, and healthcare systems, emphasizing the urgent need for collective efforts to reduce the impact of AD. Supporting and advancing research in Alzheimer's empowers individuals to make informed choices for their well-being, contributing to a future where the burden of this neurodegenerative disease is lightened.

Our initial plan involved downloading a set of raw MRI images from ADNI. We intended to preprocess these images using Clinica, applying filters such as GradWarp, B1 Correction, N3, and Scaling. However, Heudiconv, a file converter tool used before Clinica, did not successfully prepare the MRI scans for preprocessing. Consequently, we discarded Clinica and downloaded a new set of ADNI's preprocessed scans instead. Moving forward, we will show our results with Clinica, extract volumetric data from these scans, and apply machine learning algorithms to identify the brain regions associated with Alzheimer's disease. This report documents the outcomes of this research endeavor.

In addition to our study, other research has explored diverse ML methodologies and datasets to advance AD diagnosis and understanding. In [61], researchers devised an ML workflow to interpret black-box models using model-agnostic Shapley values. Unlike traditional feature importance techniques, this approach provided individual explanations for each subject and examined the intricate relationships between features and predictions. The effectiveness of the workflow was assessed by training XGBoost and RF models to analyze various stages of AD. The ADNI and the Australian Imaging, Biomarker, and Lifestyle flagship study of Ageing (AIBL) cohorts were used for model training, totaling 1700 scans. Another study proposed a novel method for simultaneous differential diagnosis of CN, MCI, and AD by combining volumetric measurements, cortical thickness measurements, hippocampal

texture, and hippocampal shape from structural MRI scans [62]. This approach uses linear discriminant analysis (LDA) classification to integrate multiple MRI biomarkers. The method was trained on 504 ADNI subjects and AIBL data, achieving comparable classification accuracy and area under the receiver operating characteristic curve to the Computer-Aided Diagnosis of Dementia (CADDementia) challenge. In [63], the deep learning-based FastSurfer pipeline was compared to the FreeSurfer pipeline for extracting volumetric features from MRI scans. FastSurfer demonstrated remarkably shorter execution times than FreeSurfer. The authors trained RF models on data from the ADNI cohort. They validated them on an independent test set within ADNI and a subset of the AIBL dataset, comprising 1565 scans from ADNI and 545 scans from AIBL. Results showed similar performance between models trained on FastSurfer and FreeSurfer data. The study in [64] introduces an image-fusion technique for integrating PET and T1-weighted MRI scans, enhancing feature fusion for AD detection. Additionally, ensemble classification methods, including Gradient Boosting (GB) and Support Vector Machine with Radial Basis Function (SVM_RBF) for Multi-Class classification and SVM_RBF + AdaBoost + GB + RF for Binary-Class classification, were proposed. The analysis included 600 scans from ADNI.

The study in [65] developed an automated ML method for classifying different stages of cognitive impairment, using cortical thickness measurements from 1167 MRI scans and achieving an overall accuracy of 75%. The authors in [66] presented a multi-modality classification framework using RF classifiers, integrating MRI volumes, FDG-PET signals, CSF biomarkers, and genetic data to achieve classification accuracies of 89% for AD vs. HC and 75% for MCI vs. HC. In [67], the authors presented the efficacy of RF classifiers using structural MRI measures, achieving a sensitivity/specificity of 88.6%/92.0% in distinguishing AD from HC, with im-

proved sensitivity/specificity for predicting MCI-to-AD conversion when including demographic and genetic data. The study in [68] focused on predicting MCI-to-AD conversion using an RF model trained on clinical data from 383 early MCI patients, resulting in an accuracy of 93.6%. Finally, the authors in [69] tested RF models on 2250 MRI scans, achieving high performance (90.2%), and identified the hippocampus, amygdala, and inferior lateral ventricle as major contributors to classification accuracy.

Our study builds upon these methodologies by integrating the Random Forest Algorithm with various cross-validation techniques to predict Alzheimer's disease and detect neuroanatomical markers of AD. We achieved an accuracy of 92.87% and an F1 score of 92.84% on our dataset, demonstrating robust predictive performance. These diverse approaches highlight the ongoing advancements in ML methodologies and their application to neuroimaging data, demonstrating the potential to substantially enhance AD diagnosis and understanding.

In this report, we use the ADNI dataset to obtain brain MRI scans. We then parcellate the neuroimages into various brain regions to create a dataset for the ML methods. The brain regions crucial in predicting AD across several subgroups are highlighted and discussed. We have covered the following topics. Section 2 shows the methodology. Section 3 presents the results and analysis of the study. Section 4 discusses and concludes the paper with our overall findings.

Chapter 2

Methodology

2.1 ADNI Dataset

The dataset used in our study includes 815 structural MRI scans (281 Cognitive Normal (CN), 332 Mild Cognitive Impairment (MCI), and 202 Alzheimer's disease (AD)) from 344 subjects, all aged between 69 and 84 years. These samples underwent preprocessing techniques such as Multiplanar Reconstruction (MPR), GradWarp, B1 Correction, N3, and Scaling. We used T1-weighted Magnetization-Prepared Rapid Gradient-Echo (MPRAGE) scans from ADNI-1, the first phase of Alzheimer's Disease Neuroimaging Initiative (ADNI).

Furthermore, subjects took the Mini-Mental State Exam (MMSE) [70] and Clinical Dementia Rating (CDR) [71] tests. The MMSE and CDR are typically administered around the time of imaging scans. The MMSE provides a quick evaluation of cognitive function, covering domains such as orientation to time, orientation to place, registration, attention and calculation, recall, language, repetition, and the ability to follow complex commands and write a sentence. The CDR relies on informant reports to stage the overall severity of dementia. These tools are essential

for Alzheimer's research as they offer standardized and quantifiable measures of cognitive decline over time. We considered the scores from these assessments when selecting subjects for our dataset. Table 2.1 provides an overview of the demographic characteristics of the ADNI dataset used in this study.

Table 2.1: ADNI demographic characteristics

	N	Sex (F:M)	Age (years)	MMSE	CDR	Modality	Image Description
CN	281	171:110	73.7-81.0	29-30	0	T1	MPR; GradWarp; B1 Correction; N3; Scaled
MCI	332	104:228	71.6-83.0	25.4-29	0.5	T1	MPR; GradWarp; B1 Correction; N3; Scaled
AD	202	108:94	69.4-83.6	16.0-25.0	0.5-1.0	T1	MPR; GradWarp; B1 Correction; N3; Scaled
	815	**383:432**	**69.4-83.6**			**T1**	**MPR; GradWarp; B1 Correction; N3; Scaled**

We categorized the dataset into distinct subgroups based on sex and age. These subgroups encompassed male and female subjects across two age ranges: 69 to 76 years and 77 to 84. Specifically, the subgroups included:

1. Both male and female subjects aged 69 to 84
2. Male-only subjects aged 69 to 84
3. Female-only subjects aged 69 to 84
4. Both male and female subjects aged 69 to 76
5. Male-only subjects aged 69 to 76
6. Female-only subjects aged 69 to 76
7. Both male and female subjects aged 77 to 84
8. Male-only subjects aged 77 to 84
9. Female-only subjects aged 77 to 84

The distribution of diagnoses and the number of samples within each of these subgroups are illustrated in Table 2.2.

Table 2.2: Distribution of ADNI Participants by Age, Sex, and Diagnosis

Age Group (in years)	**Group**	**N** **Both sexes** (M+F)	**Males** (M)	**Females** (F)
Ages 69-84	CN	281	110	171
	MCI	332	228	104
	AD	202	94	108
		815	**432**	**383**
Ages 69-76	CN	117	51	66
	MCI	151	112	39
	AD	106	39	67
		374	**202**	**172**
Ages 77-84	CN	164	59	105
	MCI	181	116	65
	AD	96	55	41
		441	**230**	**211**

2.2 Preprocessing

Preprocessing is crucial because it addresses challenges in MRI data, like variations in image quality, artifacts, and anatomical differences between subjects. Through standardization and data cleaning, preprocessing optimizes the input for deep learning and ML algorithms, making model training more accurate and reliable. This step ensures that algorithms concentrate on relevant image patterns and features, improving the model's ability to generalize and make meaningful predictions. In turn, this enhances the overall performance of neuroimaging analyses.

2.2.1 Clinica

Preprocessing in neuroimaging aims to enhance the quality and suitability of the images, addressing various issues that may arise during acquisition. Common preprocessing steps include normalization, where the intensities of the images are standardized to a consistent scale, and spatial normalization, aligning images to a standard anatomical template. Additionally, noise reduction, removal of artifacts, and correc-

tion of image distortions are essential aspects of preprocessing. A few tools required for preprocessing include:

1. Clinica [72] powers neuroimaging research with advanced pipelines, automating data preprocessing and analysis for streamlined workflows.

2. Clinica leverages the BIDS (Brain Imaging Data Structure) [73] standard, ensuring seamless brain imaging and behavioral data processing. BIDS's standardized organization simplifies data handling and improves collaboration by enabling researchers to share and utilize datasets efficiently.

3. Heudiconv [74] is a flexible DICOM converter for organizing brain imaging data into structured directory layouts. It allows flexible directory layouts and naming schemes through customizable heuristics implementations, and it only converts the necessary DICOMs and ignores everything else in a directory. It is a fast converter using dcm2niix [75] under the hood.

The initial dataset was downloaded from the ADNI server in DICOM format. This dataset was then converted into the BIDS format (see Figure 2.1) to ensure compatibility with the Clinica software. DICOM files were converted to Nifti files within the BIDS architecture. This conversion standardized the data structure, facilitating efficient analysis.

The preprocessing on Clinica began with processing the BIDS-compliant Nifti files to address noise and unwanted factors. A BIDS-compliant directory structure was created to house the converted Nifti files, ensuring proper organization and accessibility of the data. HeudiConv, a tool for DICOM-to-BIDS conversion, was used to populate the BIDS directory with Nifti files. This tool automatically [76] extracted relevant information from DICOM headers and used dcm2niix for efficient conversion. Other potential methods for this conversion included BIDScoin [77],

Bidsify [78], Bidskit [79], Data2Bids [80], Dcm2bids [81], niix2bids [82], dac2bids [83]. Specific configuration options were provided to HeudiConv to ensure accurate conversion and adherence to BIDS standards, such as specifying the output format (e.g., Nifti-1) and naming conventions for the generated files.
The preprocessing steps involved four specific tasks:

1. **GradWarp:** Corrected gradient warping artifacts in the data.
2. **B1 Correction:** Normalized intensity variations caused by radiofrequency inhomogeneities.
3. **N3 Bias Field Correction:** Removed non-uniform background intensity variations.
4. **Scaling:** Adjusted the data to a standard intensity range for improved analysis and visualization.

Finally, the preprocessed data was available in Nifti format within the BIDS structure. This processed data was free from noise and other artifacts, providing a more reliable basis for subsequent analysis.

Preprocessing software packages were used to prepare data for further ML analysis. These packages streamline tasks such as data cleaning, transformation, and feature extraction, ensuring the dataset is appropriately formatted for training ML models. Below, we explore two widely used preprocessing software packages for this study, highlighting their functionalities and applications in creating ML-ready datasets.

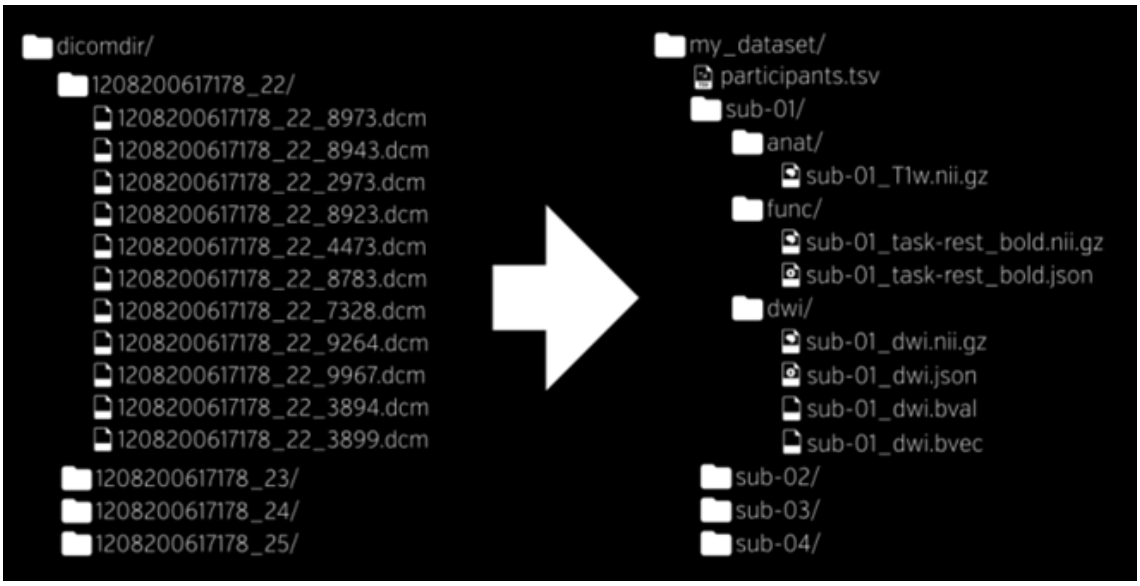


Figure 2.1: Illustration of the conversion to BIDS architecture. Each subject's directory begins with 'sub-', followed by the type of MRI scan and its corresponding .nii files. The 'anat' folder contains the anatomical (sMRI) scans, 'func' stores the functional (fMRI) scans, and 'dwi' holds diffusion-weighted imaging files, which measure the random Brownian motion of water molecules within a tissue voxel [84]. Source: `https://bids.neuroimaging.io/`.

2.2.2 FreeSurfer

FreeSurfer is an open-source neuroimaging data analysis package [85] that offers comprehensive processing capabilities, including skull-stripping, bias field correction, and cortical surface reconstruction.

We chose FreeSurfer for its robust Brain MRI analysis capabilities in our study on AD, MCI, and CN groups. Key functionalities such as:

- *Detailed Segmentation*: Differentiating gray matter, white matter, and CSF provided insights into potential neuroanatomical variations.

- *Cortical Thickness Analysis*: Enabled measurement and comparison of cortical thickness across groups, potentially revealing patterns linked to AD/MCI progression.

- *Volume Measurements*: Allowed assessment of potential volume differences in brain regions affected by AD, aiding prediction efforts.

We employed FreeSurfer's `recon-all` command to thoroughly analyze a brain MRI scan from a CN subject. This command performs a comprehensive analysis that simultaneously calculates all brain regions' volumes, thickness, and surface area. FreeSurfer's `recon-all` command is designed to process these three types of data together, and it does not allow for the individual processing of just volumes, thickness, or surface area data in isolation. The analysis was performed on two separate machines: a personal laptop and a lab machine. We successfully obtained detailed segmentation, brain region volumes, and cortical thickness.

2.2.3 FastSurfer

FastSurfer is a rapid and precise neuroimaging pipeline based on deep learning [86]. It is a swift and fully compatible alternative to FreeSurfer for volumetric and surface-based thickness analyses.

The FastSurfer pipeline comprises two primary components: segmentation and surface reconstruction. The segmentation sub-pipeline (seg) utilizes advanced deep learning networks to swiftly and accurately segment and calculate volumes for the entire Brain and specific substructures. Meanwhile, the surface sub-pipeline (recon-surf) reconstructs cortical surfaces, assigns cortical labels, and conducts traditional point-wise and ROI thickness analyses.

FastSurfer requires high-quality MRI images as input similar to those expected by FreeSurfer for optimal performance. In our study, we used FastSurfer's `--seg_only` and `--surf_only` commands to respectively analyze brain volume and thickness from a single MRI scan of a CN subject. This analysis was performed on the personal machine and the lab machine. However, when using the `--surf_only` command,

we encountered difficulty parcellating the cerebellum on both machines. Since the literature indicates that the cerebellum is unaffected by AD [87], we decided to exclude it from our analysis using the `--no_cereb` flag and continued our research. Table 2.3 provides a comprehensive comparison between FreeSurfer and FastSurfer pipelines.

Table 2.4 shows the specifications of the machines and their processing times for

Table 2.3: Comparison of FreeSurfer and FastSurfer

Feature	**FreeSurfer**	**FastSurfer**
Type	Traditional neuroimaging analysis pipeline	Deep learning-based neuroimaging pipeline
Accuracy	Well-established, high accuracy	High accuracy, comparable to FreeSurfer
Speed	Relatively slow, computationally expensive	Significantly faster than FreeSurfer
Applications	Research studies, clinical applications	Research studies (faster analysis), screening applications

one scan using FreeSurfer and FastSurfer pipelines. Both analyses yielded promising results. However, both machines' high overall processing time for the FreeSurfer pipeline led us to discontinue using it for further studies. The processing time with the fastest settings for volumetric analysis using FastSurfer was 1 minute per scan, while thickness analysis took approximately 24 minutes. These processing times prominently outpaced FreeSurfer's 3-hour processing time per scan. Thus, we opted to continue our research using FastSurfer. To decide between volumetric and surface-based thickness analysis, we focused on the processing time required by the `--seg_only` and `--surf_only` commands for segmenting the entire dataset, as seen in Table 2.5. We initiated our study with volumetric analysis as it required less processing time for the 815 scans.

In this study, we focused on measuring volume changes in both cortical and subcortical brain regions. This approach was chosen to maintain consistency in our analyses. All regions, i.e., cortical and subcortical areas, were analyzed for volume

Table 2.4: Comparison of FreeSurfer and FastSurfer Analyses on Different Machines

Machine	**Operating System**	**Hardware**	**Processing Time (per scan)**	
			FreeSurfer	**FastSurfer**
Personal Machine	Ubuntu 22.04 (VMware Workstation 17 Player)	Laptop with 32GB RAM and 4GB graphics card	Volumetric & thickness analysis ≈ 12 hours	Volumetric analysis ≈ 30 minutes Thickness analysis ≈ 3 hours
Lab Machine	Ubuntu 22.04	Computer with 64GB RAM, and NVIDIA GeForce RTX 2080 SUPER Graphics Processor, with 8GB graphics memory	Volumetric & thickness analysis ≈ 3 hours	Volumetric analysis ≈ 1 minute Thickness analysis ≈ 24 minutes

Table 2.5: Processing Time for Volumetric and Thickness Analysis

Analysis Type	**Processing Time (for 815 scans)**
Volumetric Analysis (`--seg_only`)	≈ 13.5 hours
Thickness Analysis (`--surf_only`)	≈ 13+ days

changes to identify structural differences associated with age, sex, and AD.

2.3 Class Imbalance

Class imbalance refers to the scenario where the distribution of classes in the dataset is uneven, with one class considerably outnumbering the others. In ML tasks, class imbalance can pose challenges as models favor the majority class, leading to poor performance in minority classes. Various techniques are available to address these challenges during data preprocessing and model training, such as using oversampling or undersampling methods to balance the class distribution.

In this study, we employ the Synthetic Minority Over-sampling Technique (SMOTE) from the `imblearn` library to oversample the minority class [88]. SMOTE generates synthetic samples from the minority class by interpolating new instances between existing minority class samples. This technique helps to alleviate class imbalance by creating more balanced datasets for training machine learning models, thereby improving their performance, especially in predicting minority class instances. Alternative methods, such as ADASYN [89], can be utilized; however, we opted not to employ ADASYN in this study as the evidence demonstrates that SMOTE outperforms ADASYN in various scenarios [90].

Class imbalance occurs after preprocessing the dataset with normalization techniques and splitting it into training and testing sets. When training random forest classifiers, we apply the `SMOTE` oversampling technique to the training data to ensure a balanced class distribution. The impact of class imbalance on model performance affects the analysis of top contributing features and performance metrics. By addressing the class imbalance, we aim to improve the robustness of the RF classifiers for predicting AD. Notably, the accuracy obtained without using SMOTE was in the range 86-88%, but it improved to the range 90-93% with SMOTE.

2.4 Classification Model: Random Forest

Random Forest (RF) is a popular ensemble learning algorithm widely used for classification and regression tasks in machine learning [91]. It operates by constructing multiple decision trees during training. Each tree is trained on a bootstrapped dataset sample, and a random subset of features is considered for splitting at each tree node. The final prediction is then made by aggregating the predictions of all individual trees in the ensemble. RF can thus output the mode of the classes

(classification) and the individual trees' mean prediction (regression). RF can also generate a list of ranked features by measuring the importance of each feature in reducing impurity or error within the forest's trees. This ranking is based on how much each feature contributes to the model's overall accuracy. In this study, we have implemented the Random Forest algorithm using the scikit-learn library in Python. The flowchart depicted in Figure 2.2 provides an example of how RF predicts Alzheimer's Disease.

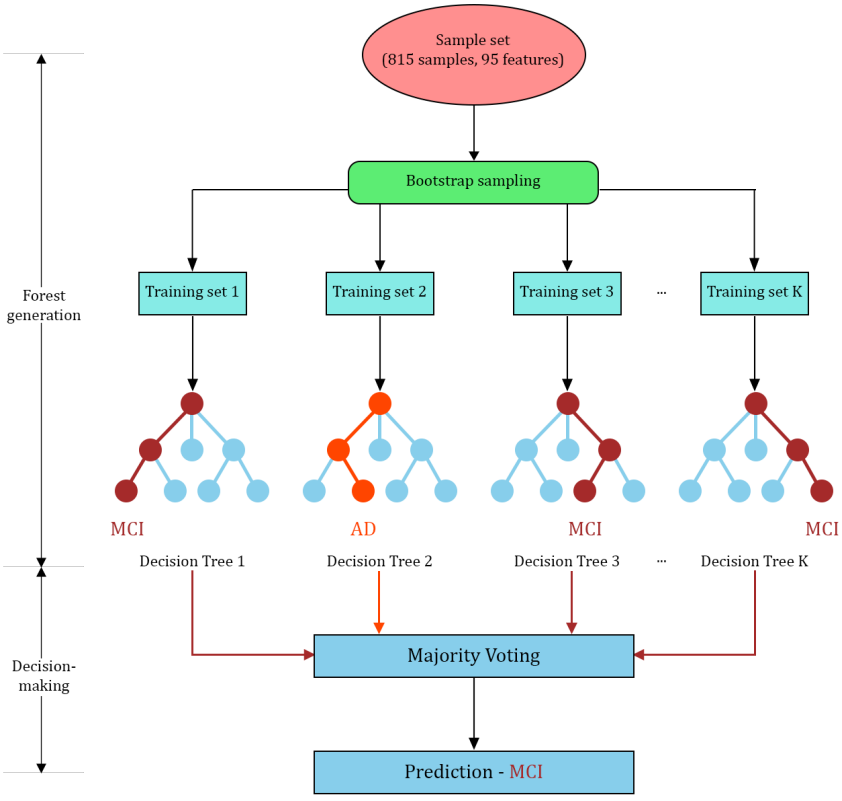


Figure 2.2: An example flowchart of Random Forest classifier.

We extend the use of the ML algorithm in [92] to predict AD. The primary objective is to identify the brain regions driving AD from MRI scans. Here's how the RF algorithm is implemented:

1. **Number of Trees (Estimators):**

    - The RF classifier is configured with varying numbers of trees, also known as estimators, representing the number of decision trees to be constructed. In the hyperparameter tuning process, the number of estimators is explored across different values, such as 100, 200, and 300, to determine the optimal setting. The optimal number of trees is typically selected based on cross-validation performance metrics, such as accuracy or area under the ROC curve (AUC), where the model achieves the highest performance without overfitting the training data.
    - Increasing the estimators in RF generally leads to better performance up to a certain point. More trees can help capture complex relationships in the data and reduce overfitting.
    - However, using too many trees may increase computational complexity and training time.

2. **Maximum Depth of Trees:**

    - The maximum depth of each tree in the forest influences the complexity and generalization capability of the model.
    - During hyperparameter tuning, different values for the maximum depth, including `None`, 10, 20, and 30, were explored to identify the optimal setting.
    - This parameter controls the level of detail in the model. Deeper trees can capture more intricate patterns in the data but may increase the risk of overfitting.
    - It's important to note that increasing the maximum depth beyond a

certain threshold can lead to overfitting, where the model memorizes noise in the training data rather than learning generalizable patterns.

3. **Minimum Samples Split and Minimum Samples Leaf:**

   - Different combinations of hyperparameters related to node splitting are explored, including the minimum number of samples required to split an internal node (`min_samples_split`) and the minimum number of samples required to be a leaf node (`min_samples_leaf`). This exploration aims to optimize the decision tree construction process and enhance the classifier's performance.
   - Increasing the values of `min_samples_split` and `min_samples_leaf` can help regularize the model and prevent overfitting by ensuring that each node contains a sufficient amount of data for splitting or becoming a leaf.
   - However, setting excessively high values for these hyperparameters may result in underfitting, where the model fails to capture important patterns in the data due to overly strict splitting criteria.

4. **Model Training:**

   - After identifying the best hyperparameters through hyperparameter tuning, the RF classifier is trained with this optimized configuration. This process produces a finely tuned model qualified for making accurate predictions.
   - It is important to monitor the model's performance on separate validation or test data sets. This step ensures that the model's effectiveness extends beyond the training data and can generalize well to new, unseen data instances.

2.5 RF-Based Neuroanatomical Marker Analysis

One of the critical advantages of RF is its ability to identify the most informative features for classification tasks. This study explores how RF determines the top contributing features and assesses their importance.

2.5.1 Understanding Gini Impurity

In the context of decision trees, Gini impurity measures how often a randomly chosen element from the set would be incorrectly labeled if it was randomly labeled according to the distribution of labels in the set. It is determined by subtracting the sum of the squares of the probabilities of each class from one. The Gini impurity for a node t in a decision tree is calculated as follows:

$$\text{Gini}(t) = 1 - \sum_{i=1}^{C} p(i|t)^2$$

Where, C is the number of classes, and
$p(i|t)$ is the probability of class i at node t.

In technical terms, The Gini impurity measures the probability of misclassifying a randomly chosen sample's label, weighted by the probability of selecting that label in the node. A Gini impurity of 0 indicates perfect purity (all samples belong to the same class), while a Gini impurity of 1 indicates maximum impurity (samples are evenly distributed across all classes).

2.5.2 Feature Importance

In RF, the importance of a feature is determined by measuring the weighted average of the decrease in Gini impurity resulting from splitting on that feature across

all trees in the ensemble. Impurity, often calculated using Gini impurity or entropy, quantifies the uncertainty in class labels at a given node. The decrease in impurity resulting from splitting on a particular feature is a measure of its importance. Features with higher decreases in impurity are considered more important for classification. The importance of a feature X_i is calculated as:

$$\text{Importance}(X_i) = \frac{\sum_{t \in \text{Trees}} \text{Gini}(t) \times \text{Number of Samples in Node Split on } X_i}{\sum_{t \in \text{Trees}} \text{Number of Samples in Node Split on } X_i}$$

Where:

Trees is the set of all decision trees in the Random Forest, and

Gini(t) is the Gini impurity of the node split on feature X_i in tree t.

After training the RF model, we extract the feature importance scores provided by the model. These scores represent the relative importance of each feature in contributing to the classification task. We then rank the features based on their importance scores to identify the top contributing features for predicting AD.

2.6 Model Validation

In this section, we discuss the validation techniques employed to assess the performance of the RF models for predicting AD.

2.6.1 K-Fold Cross-Validation

K-fold cross-validation divides the dataset into k equal-sized folds, where each fold acts as a validation set while the remaining $k-1$ folds are used for training. This process is repeated k times, with each fold serving as the validation set precisely once. We evaluate the model's performance by averaging the results across all folds.

In this study, we put $k = 5$, splitting the dataset into five folds for cross-validation.

The trend observed in K-Fold Cross-Validation is that the model's performance metrics, such as accuracy, precision, recall, and F1-score, tend to stabilize as the number of folds increases. A higher number of folds generally leads to a more reliable estimate of the model's performance.

2.6.2 Stratified K-Fold Cross-Validation

Stratified K-Fold Cross-Validation preserves the class distribution in each fold, ensuring that each fold is representative of the overall dataset's class distribution. This technique benefits imbalanced datasets, where one class may be considerably smaller than the others. We set $k = 5$ for Stratified K-Fold Cross-Validation in this study.

Stratified K-Fold Cross-Validation exhibits a similar trend to K-Fold Cross-Validation, with the added advantage of ensuring that each fold maintains the class balance observed in the entire dataset.

2.6.3 Leave-One-Out Cross-Validation

Leave-one-out cross-validation (LOOCV) is a technique where a single sample is used as the validation set, and the rest of the samples are used for training [93]. This process is repeated for each sample in the dataset, resulting in n folds where n is the number of samples. In this study, we employed LOOCV to evaluate the performance of our models.

In Leave-One-Out Cross-Validation (LOOCV), we observe a trend where the model's performance metrics, such as accuracy, precision, recall, and F1-score, tend to stabilize as the size of the dataset increases. With LOOCV, the model is repeatedly trained and evaluated on nearly identical datasets, resulting in a comprehensive

assessment of its performance.

2.6.4 Implementation

In our implementation, we use the `cross_val_predict` function from the `sklearn.model_selection` module to perform K-Fold, Stratified K-Fold, and Leave-one-out cross-validation. This function efficiently handles the splitting of the dataset and evaluation of the model across multiple folds.

2.7 Performance Metrics

Performance metrics are essential tools for evaluating the effectiveness and efficiency of machine learning models [94]. These metrics provide insights into how well the model performs in terms of accuracy, precision, recall, F1-score, and other relevant measures. The following are the various performance metrics used in this study to evaluate the RF classifier's effectiveness in predicting AD:

1. **Accuracy:**

    - Accuracy measures the proportion of correctly classified instances out of the total instances in the dataset. It is calculated as the ratio of the number of correct predictions to the total number of predictions.
    - It performs well when classes are balanced, and no considerable class imbalance exists. However, accuracy can be misleading in imbalanced datasets where the majority class dominates.

2. **Precision:**

    - Precision quantifies the model's ability to correctly identify positive instances (true positives) out of all instances classified as positive. It is

calculated as the ratio of true positives to the sum of true positives and false positives.

- Precision may be high when the model is conservative in predicting positive instances but may suffer when the model misses some true positives, resulting in a low recall.

3. **Recall:**

   - Recall, also known as sensitivity, measures the model's ability to identify positive instances out of all actual positive instances correctly. It is calculated as the ratio of true positives to the sum of true positives and false negatives.
   - High recall may lead to more false positives, which can be problematic when false alarms have severe consequences, such as medical diagnoses or fraud detection.

4. **F1-score:**

   - The F1-score represents the harmonic mean of precision and recall, offering a balanced assessment between the two metrics. It proves especially valuable when both precision and recall are important, particularly in scenarios of an imbalance between classes. By providing a unified metric, it effectively balances the consideration of false positives and false negatives.

These performance metrics are computed and reported using functions such as `accuracy_score`, `precision_score`, `recall_score`, and `f1_score` from the `sklearn.metrics` module.

In summary, our methodology involved using the ADNI dataset for MRI scans, attempting to preprocess the scans using Clinica, and conducting volumetric analysis using FastSurfer. We initially categorized the dataset into male and female cohorts to investigate sex-specific disparities in AD progression. We then divided these cohorts into two age groups to examine age-specific differences. We employed an RF Classifier, SMOTE, and three validation techniques to pinpoint the brain regions influencing AD. Performance metrics were assessed, and the next chapter presents the obtained results.

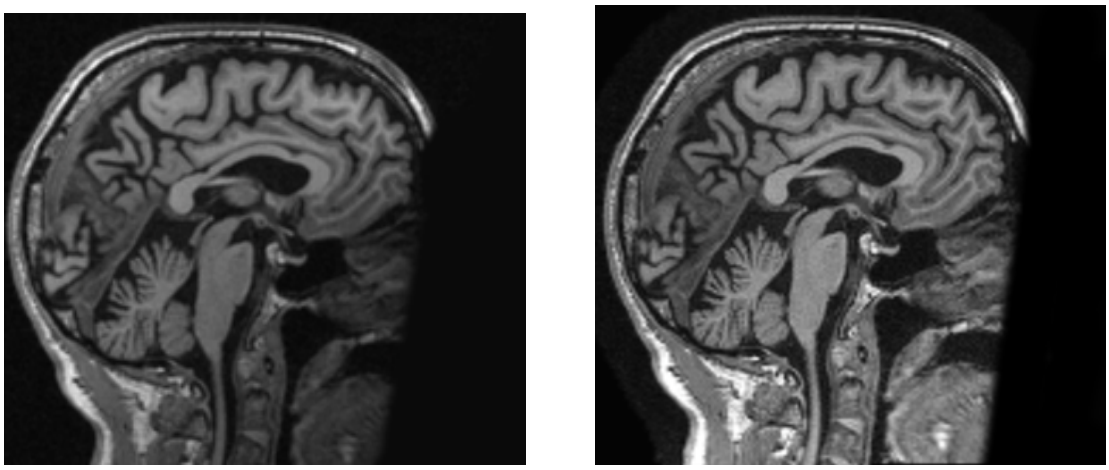
Chapter 3

Results

3.1 Clinica outputs

Figure 3.1 shows two scans of the same patient before and after preprocessing with Clinica software. The workflow began with converting the raw DICOM MRI scans from the ADNI dataset to BIDS format. Next, the HeudiConv tool converted these DICOM files to Nifti format. Following these conversions, Clinica was used to apply filters such as GradWarp, B1 Correction, N3 Bias Field Correction, and Scaling to the Nifti brain scans. The resulting output is depicted in Figure 3.1(b). Moreover, standard practices like skull stripping and cropping could be employed to focus on critical structures during neuroimaging analysis.

The use of HeudiConv resulted in substantial data loss during the conversion process, with only about 30 out of 350 files successfully converted to Nifti format. This conversion success rate could potentially be improved with further expertise in the tool. However, due to the poor results from Heudiconv, we downloaded preprocessed images directly from the ADNI dataset. This approach ensured better data integrity and allowed for more reliable subsequent analyses.



(a) (b)

Figure 3.1: Defaced sMRI scans (a) Original (b) Preprocessed.

3.2 FastSurfer outputs

The FastSurfer's `--seg-only` pipeline took preprocessed MRI scans from ADNI, in Nifti (.nii) file format, as input and gave MRI images parcellated into 100 brain regions. These regions consisted of 31 cortical areas per hemisphere, 33 subcortical regions, and five areas of the corpus callosum, as defined by the Desikan-Killiany-Tourville (DKT) Atlas, which is the standard atlas* utilized by FastSurfer. Additionally, FastSurfer provided volumetric information for each region in a text file for every scan.

Out of the 100 brain regions obtained, we excluded those with zero volumes, resulting in 95 brain regions being considered. A list of these 95 brain regions is given in the Supplementary section in Chapter 5. Figure 3.2 visualizes a CN subject's brain parcellation as seen in Freeview neuroimaging software.

*An atlas serves as a comprehensive reference for understanding the spatial organization of the brain and its structural connectivity [95]. It facilitates the precise localization of specific brain regions and the analysis of their functional activities and enables comparisons with standardized datasets. By integrating various perspectives of the brain and employing registration techniques, atlases offer a systematic framework for studying brain anatomy and function, thereby enhancing our understanding of its complex architecture.

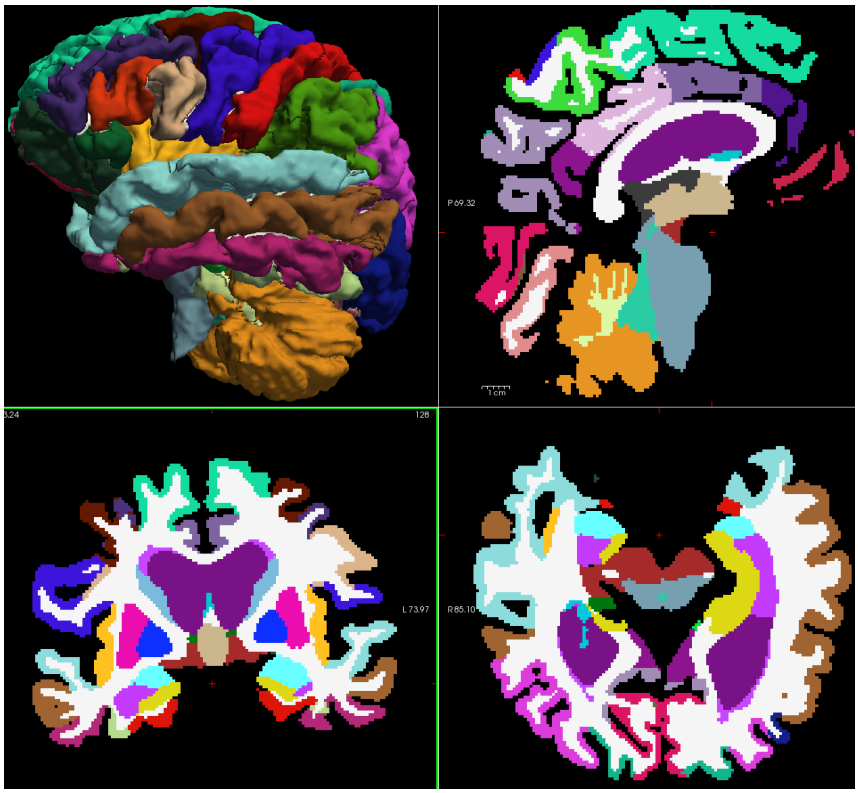


Figure 3.2: Brain parcellation in Freeview.

Using the DKT atlas alongside the `--seg-only` pipeline, only 100 brain regions were identified. Alternatively, employing the `--surf-only` pipeline would have yielded a more extensive set comprising 172 brain regions. This expanded set would have included additional regions such as white matter regions, both hemispheres' cortex banks, cortex frontal pole, and cortex temporal pole. However, due to prioritizing computational efficiency, we opted to focus solely on the 100 brain regions generated by the `--seg-only` pipeline. Processing 815 scans took 1.5 hours on day 1 and 12 hours on day 2.

FreeSurfer's `aseg2stats.table` command was used to convert the above subcortical stats file into a tabular format. In this table, each line represents a subject, and each column corresponds to a segmentation, with the volume of each segmentation measured in cubic millimeters (mm^3). The first row contains the segmentation

names, and the first column lists the subject names. Demographic information of each scan was then appended to this table.

3.3 Feature Engineering

After acquiring our dataset with patient demographics, we employed feature engineering to optimize our data preprocessing strategy. Feature engineering techniques, such as normalization, are crucial as they directly impact the model's ability to learn and make accurate predictions. Despite normalization not being commonly practiced in neuroimaging studies involving human subjects, as seen from the literature, we sought guidance from Dr. Marcelo Febo, an expert in functional magnetic resonance imaging.

Dr. Febo advised against traditional normalization methods, suggesting that while normalization could address differences in brain volumes, it wasn't imperative. Instead, he recommended considering intracranial volumes as covariates in our study design. Without estimated total intracranial volumes provided by FastSurfer, we devised a novel approach. We normalized each brain region volume of a subject by the sum of all 95 regional brain volumes. This custom normalization technique ensured that each feature fell within a consistent range between 0 and 1, as illustrated in Equation (3.1):

$$\text{Normalization} = \frac{\text{Volume of a Brain Region}}{\text{Sum of Every Brain Region of That Subject}} \tag{3.1}$$

This custom normalization led to a slight improvement in data quality, reflected in our model's performance. Using RF with this custom normalization strategy, we achieved an accuracy of **0.883**, compared to **0.877** without normalization.

Moreover, traditional normalization techniques, such as `MinMaxScaler()` from

the `sklearn.preprocessing` module, were explored. This showed similar accuracy to our custom normalization, indicating comparable effectiveness in reducing confounding effects. We then pursued custom normalization further in our analysis.

Following data normalization, we organized the dataset into subgroups based on sex and age for further analysis. These subgroups encompassed both male and female subjects, male-only subjects, female-only subjects, both male and female subjects aged 69 to 76, male-only subjects aged 69 to 76, female-only subjects aged 69 to 76, both male and female subjects aged 77 to 84, male-only subjects aged 77 to 84, and female-only subjects aged 77 to 84. We applied the `SMOTE` technique to address any imbalances in class distribution. Additionally, we explored various hyperparameter search techniques, including Randomized Search and BayesSearchCV, which yielded similar accuracies. Given the slight edge in performance, we opted for the ParameterGrid approach. We then fine-tuned the Random Forest parameters through hyperparameter tuning and validated them using K-fold, stratified K-fold, and leave-one-out cross-validation methods. Finally, we identified the most influential features across all subgroups.

3.4 Subgroup Analyses

3.4.1 Statistical Analysis of Brain Regions

Several brain regions were analyzed for statistical comparisons. Figure 3.3 shows the average volumes of the left and right hippocampus, amygdala, and entorhinal cortex in CN, MCI, and AD patients. The statistical differences among the three groups were assessed using a two-sample Z-test for each pair of brain regions. Significant differences ($p < 0.01$) were found across CN vs. MCI, MCI vs. AD, and CN vs. AD comparisons. The Z-test was chosen to demonstrate significant differences in brain

region volumes, which were influenced by the consistently low p-values, emphasizing the reliability of our statistical outcomes.

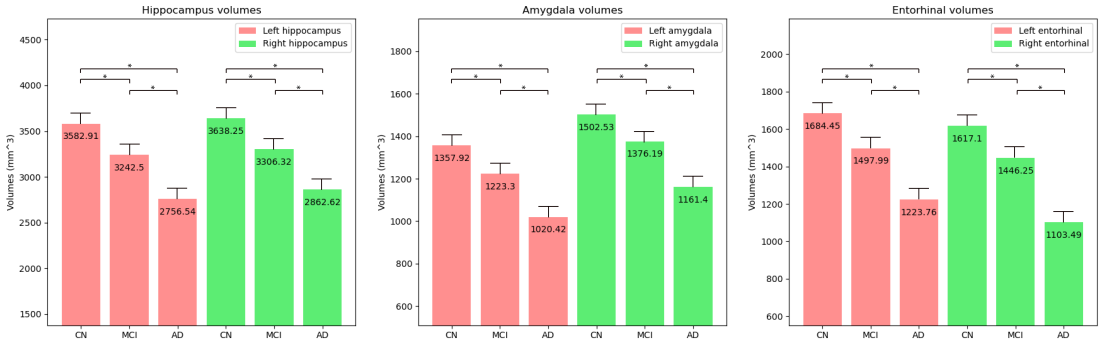


Figure 3.3: Average volumes of left and right hippocampus, amygdala, and entorhinal cortex in CN, MCI, and AD subjects, illustrating atrophy across cognitive states. Statistical significance is marked with * ($p < 0.01$).

Figure 3.4 presents the average volumes of four critical brain regions—Left-Inf-Lat-Vent (LILV), left inferior parietal cortex, left inferior temporal cortex, and left middle temporal cortex—across cognitive states in both sexes aged 69-84. Significant differences ($p < 0.01$) were observed in all comparisons except for CN to MCI in the left inferior parietal and left inferior temporal cortices, where no significant differences were found. These results highlight the impact of widespread brain atrophy in AD, with certain areas showing early volume changes and others exhibiting more pronounced differences as the disease progresses.

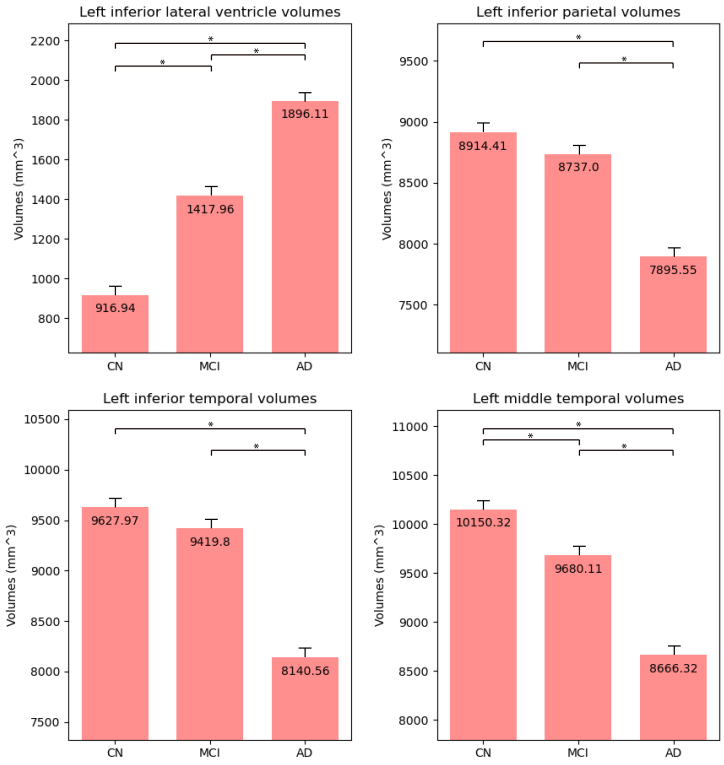


Figure 3.4: Average volumes of Left Inferior Lateral Ventricle, left inferior parietal cortex, left inferior temporal cortex, and left middle temporal cortex in CN, MCI, and AD subjects, illustrating the trend of ventricular enlargement and cortical atrophy across cognitive states. Statistical significance is marked with * ($p < 0.01$).

Understanding the Benjamini-Hochberg Correction

After conducting statistical tests to compare brain region volumes between groups (CN, MCI, and AD), we need to consider the possibility of false positives, especially when doing multiple tests. The Benjamini-Hochberg (BH) correction helps address this concern by adjusting the significance level for each comparison [96].

Before applying the BH correction, we obtained p-values from the tests. These p-values indicate the likelihood of observing the results by chance. However, they don't account for the increased risk of false positives with multiple comparisons.

After applying the BH correction, the p-values are adjusted to better control the false discovery rate. Comparing the p-values before and after the correction gives us a clearer picture of the true significance of the differences observed in brain region volumes among the groups.

Comparison of P-Values

Brain Region	CN vs MCI	CN vs AD	MCI vs AD	MCI vs CN	AD vs CN	AD vs MCI
Left-Hippocampus	5.43e-22	1.24e-92	5.43e-22	1.05e-30	1.24e-92	1.05e-30
Right-Hippocampus	6.82e-17	1.05e-68	6.82e-17	1.91e-23	1.05e-68	1.91e-23
Left-Amygdala	1.85e-13	6.97e-70	1.85e-13	9.47e-26	6.97e-70	9.47e-26
Right-Amygdala	6.16e-09	2.65e-47	6.16e-09	1.66e-19	2.65e-47	1.66e-19
ctx-lh-entorhinal	3.92e-12	3.76e-56	3.92e-12	1.41e-18	3.76e-56	1.41e-18
ctx-rh-entorhinal	3.44e-09	2.77e-69	3.44e-09	3.39e-29	2.77e-69	3.39e-29
Left-Inf-Lat-Vent	5.09e-20	1.91e-27	6.84e-07	5.09e-20	1.91e-27	6.84e-07
ctx-lh-inferiorparietal	0.136	1.73e-15	1.60e-09	0.136	1.73e-15	1.60e-09
ctx-lh-inferiortemporal	0.0793	7.32e-30	1.74e-20	0.0793	7.32e-30	1.74e-20
ctx-lh-middletemporal	2.96e-04	8.68e-24	8.28e-12	2.96e-04	8.68e-24	8.28e-12

Table 3.1: P-values before Benjamini-Hochberg correction

After Benjamini-Hochberg Correction:

Brain Region	CN vs MCI	CN vs AD	MCI vs AD	MCI vs CN	AD vs CN	AD vs MCI
Left-Hippocampus	8.89e-22	2.24e-91	8.89e-22	2.69e-30	2.24e-91	2.69e-30
Right-Hippocampus	8.77e-17	4.71e-68	8.77e-17	3.44e-23	4.71e-68	3.44e-23
Left-Amygdala	2.22e-13	6.27e-69	2.22e-13	1.89e-25	6.27e-69	1.89e-25
Right-Amygdala	6.16e-09	7.95e-47	6.16e-09	2.50e-19	7.95e-47	2.50e-19
ctx-lh-entorhinal	4.41e-12	1.36e-55	4.41e-12	1.95e-18	1.36e-55	1.95e-18
ctx-rh-entorhinal	3.64e-09	1.66e-68	3.64e-09	7.63e-29	1.66e-68	7.63e-29
Left-Inf-Lat-Vent	9.54e-20	5.72e-27	9.54e-20	7.60e-07	5.72e-27	7.60e-07
ctx-lh-inferiorparietal	0.136	2.60e-15	0.136	1.99e-09	2.60e-15	1.99e-09
ctx-lh-inferiortemporal	0.0821	2.75e-29	0.0821	3.48e-20	2.75e-29	3.48e-20
ctx-lh-middletemporal	3.17e-04	2.17e-23	3.17e-04	1.08e-11	2.17e-23	1.08e-11

Table 3.2: P-values after Benjamini-Hochberg correction

Table 3.1 shows the p-values obtained before applying the Benjamini-Hochberg correction, while Table 3.2 displays the p-values after the correction. These tables allow for comparing the significance levels before and after adjusting for multiple comparisons, providing insights into the robustness of the statistical findings. The corrected p-values from the BH correction provide a more reliable indication of the true significance of differences in brain region volumes among the groups. This adjustment helps ensure that our conclusions are based on robust statistical evidence,

enhancing the validity of our findings.

3.4.2 Performance Metrics

Tables 3.3, 3.4, and 3.5 display the aggregated performance metrics achieved through K-fold, stratified K-fold, and leave-one-out cross-validation techniques across the nine subgroups. The performance of all three validation methods is comparable, with accuracy consistently in the low 90s. Leave-one-out cross-validation, however, demonstrates slightly superior performance. The total execution time is presented, with leave-one-out validation taking the longest, at 23.73 minutes. In contrast, the other two methods complete much faster, in just 1 minute each.

Table 3.3: Comparison of Aggregated Performance Metrics using K-fold Cross-Validation

Groups	**Unified Sex (M+F)**		**Males (M)**		**Females (F)**	
Unified age group Ages 69-84	Metric	Value	Metric	Value	Metric	Value
	Accuracy	0.9086	Accuracy	0.9415	Accuracy	0.9357
	Precision	0.9084	Precision	0.9414	Precision	0.9358
	Recall	0.9086	Recall	0.9415	Recall	0.9357
	F1-score	0.9080	F1-score	0.9413	F1-score	0.9355
	Training accuracy	1.0	Training accuracy	1.0	Training accuracy	1.0
Younger age group Ages 69-76	Metric	Value	Metric	Value	Metric	Value
	Accuracy	0.9051	Accuracy	0.9315	Accuracy	0.9403
	Precision	0.9052	Precision	0.9313	Precision	0.9403
	Recall	0.9051	Recall	0.9315	Recall	0.9403
	F1-score	0.9041	F1-score	0.9312	F1-score	0.9403
	Training accuracy	1.0	Training accuracy	1.0	Training accuracy	1.0
Older age group Ages 77-84	Metric	Value	Metric	Value	Metric	Value
	Accuracy	0.9098	Accuracy	0.9282	Accuracy	0.9365
	Precision	0.9095	Precision	0.9284	Precision	0.9366
	Recall	0.9098	Recall	0.9282	Recall	0.9365
	F1-score	0.9095	F1-score	0.9281	F1-score	0.9360
	Training accuracy	1.0	Training accuracy	1.0	Training accuracy	1.0
Total execution time for K-fold: **1.31 minutes** (78.76 seconds)						

3.4.3 RF-Based Neuroanatomical Marker Analysis

The tables 3.6, 3.7, and 3.8 present the top six brain regions implicated in Alzheimer's Disease progression across nine subgroups, identified through K-fold, stratified K-fold, and leave-one-out cross-validation techniques. These tables reveal consistent

Table 3.4: Comparison of Aggregated Performance Metrics using Stratified K-fold Cross-Validation

Groups	Unified Sex (M+F)		Males (M)		Females (F)	
Unified age group Ages 69-84	Metric	Value	Metric	Value	Metric	Value
	Accuracy	0.9187	Accuracy	0.9503	Accuracy	0.9357
	Precision	0.9190	Precision	0.9502	Precision	0.9359
	Recall	0.9187	Recall	0.9503	Recall	0.9357
	F1-score	0.9178	F1-score	0.9502	F1-score	0.9355
	Training accuracy	1.0	Training accuracy	1.0	Training accuracy	1.0
Younger age group Ages 69-76	Metric	Value	Metric	Value	Metric	Value
	Accuracy	0.8962	Accuracy	0.9583	Accuracy	0.9353
	Precision	0.8962	Precision	0.9583	Precision	0.9355
	Recall	0.8962	Recall	0.9583	Recall	0.9353
	F1-score	0.8957	F1-score	0.9582	F1-score	0.9351
	Training accuracy	1.0	Training accuracy	1.0	Training accuracy	1.0
Older age group Ages 77-84	Metric	Value	Metric	Value	Metric	Value
	Accuracy	0.9116	Accuracy	0.9224	Accuracy	0.9365
	Precision	0.9123	Precision	0.9222	Precision	0.9363
	Recall	0.9116	Recall	0.9224	Recall	0.9365
	F1-score	0.9107	F1-score	0.9222	F1-score	0.9363
	Training accuracy	1.0	Training accuracy	1.0	Training accuracy	1.0
Total execution time for Stratified K-fold: **1.34 minutes** (80.67 seconds)						

Table 3.5: Comparison of Aggregated Performance Metrics using Leave-One-Out Cross-Validation

Groups	Unified Sex (M+F)		Males (M)		Females (F)	
Unified age group Ages 69-84	Metric	Value	Metric	Value	Metric	Value
	Accuracy	0.9287	Accuracy	0.9547	Accuracy	0.9435
	Precision	0.9284	Precision	0.9547	Precision	0.9435
	Recall	0.9287	Recall	0.9547	Recall	0.9435
	F1-score	0.9284	F1-score	0.9546	F1-score	0.9434
	Training accuracy	1.0	Training accuracy	1.0	Training accuracy	1.0
Younger age group Ages 69-76	Metric	Value	Metric	Value	Metric	Value
	Accuracy	0.9205	Accuracy	0.9435	Accuracy	0.9303
	Precision	0.9212	Precision	0.9435	Precision	0.9307
	Recall	0.9205	Recall	0.9435	Recall	0.9303
	F1-score	0.9200	F1-score	0.9430	F1-score	0.9304
	Training accuracy	1.0	Training accuracy	1.0	Training accuracy	1.0
Older age group Ages 77-84	Metric	Value	Metric	Value	Metric	Value
	Accuracy	0.9263	Accuracy	0.9397	Accuracy	0.9460
	Precision	0.9272	Precision	0.9396	Precision	0.9477
	Recall	0.9263	Recall	0.9397	Recall	0.9460
	F1-score	0.9257	F1-score	0.9396	F1-score	0.9455
	Training accuracy	1.0	Training accuracy	1.0	Training accuracy	1.0
Total execution time for Leave-One-Out: **23.73 minutes**						

brain regions recurring as top contributors across the subgroups.

Table 3.6: Comparison of Top Contributing Features using K-fold Cross-Validation

Groups	Unified Sex (M+F)		Males (M)		Females (F)	
	Feature	Importance	Feature	Importance	Feature	Importance
Unified age group Ages 69-84	Left-Hippocampus	0.0592	Left-Hippocampus	0.04999	Left-Hippocampus	0.0694
	Left-Amygdala	0.0406	Left-Amygdala	0.0381	Left-Amygdala	0.0470
	Right-Amygdala	0.0341	ctx-rh-entorhinal	0.0355	Right-Hippocampus	0.0352
	ctx-rh-entorhinal	0.0299	ctx-lh-inferiorparietal	0.0333	ctx-lh-entorhinal	0.0334
	Right-Hippocampus	0.0296	Right-Hippocampus	0.0323	ctx-lh-inferiortemporal	0.0266
	ctx-lh-inferiortemporal	0.0251	ctx-lh-middletemporal	0.0289	ctx-lh-middletemporal	0.0261
	Feature	Importance	Feature	Importance	Feature	Importance
Younger age group Ages 69-76	Left-Hippocampus	0.0565	Right-Hippocampus	0.0603	Left-Hippocampus	0.0674
	Right-Hippocampus	0.0356	Left-Hippocampus	0.0443	Left-Amygdala	0.0434
	Left-Amygdala	0.0339	Right-Amygdala	0.0423	ctx-lh-parahippocampal	0.0333
	ctx-lh-entorhinal	0.0298	ctx-rh-entorhinal	0.0379	ctx-lh-middletemporal	0.0319
	ctx-rh-entorhinal	0.0277	Left-Amygdala	0.0302	ctx-lh-entorhinal	0.0315
	ctx-rh-middletemporal	0.0274	ctx-lh-inferiorparietal	0.0215	ctx-lh-inferiortemporal	0.0292
	Feature	Importance	Feature	Importance	Feature	Importance
Older age group Ages 77-84	Left-Hippocampus	0.0585	Left-Hippocampus	0.0458	Left-Hippocampus	0.0796
	ctx-rh-entorhinal	0.0395	Left-Inf-Lat-Vent	0.0357	Left-Amygdala	0.0409
	Left-Amygdala	0.0378	ctx-lh-inferiortemporal	0.0323	ctx-lh-middletemporal	0.0398
	ctx-lh-entorhinal	0.0373	ctx-lh-entorhinal	0.0320	Right-Amygdala	0.0389
	ctx-lh-inferiortemporal	0.0336	ctx-rh-entorhinal	0.0316	ctx-lh-entorhinal	0.0344
	Right-Hippocampus	0.0243	Left-Amygdala	0.0308	Right-Hippocampus	0.0311

Table 3.7: Comparison of Top Contributing Features using Stratified K-fold Cross-Validation

Groups	Unified Sex (M+F)		Males (M)		Females (F)	
	Feature	Importance	Feature	Importance	Feature	Importance
Unified age group Ages 69-84	Left-Hippocampus	0.0629	Left-Hippocampus	0.0618	Left-Hippocampus	0.0708
	Left-Amygdala	0.0374	ctx-rh-entorhinal	0.0347	ctx-lh-entorhinal	0.0375
	ctx-rh-entorhinal	0.0300	Left-Amygdala	0.0293	Left-Amygdala	0.0373
	Right-Hippocampus	0.0277	Right-Amygdala	0.0286	Right-Hippocampus	0.0337
	Right-Amygdala	0.0254	ctx-lh-inferiorparietal	0.0272	ctx-lh-inferiortemporal	0.0275
	ctx-lh-inferiortemporal	0.0239	Right-Hippocampus	0.0263	Right-Amygdala	0.0248
	Feature	Importance	Feature	Importance	Feature	Importance
Younger age group Ages 69-76	Left-Hippocampus	0.0565	Right-Hippocampus	0.0488	Left-Hippocampus	0.0580
	Left-Amygdala	0.0454	Left-Hippocampus	0.0434	ctx-lh-middletemporal	0.0502
	Right-Hippocampus	0.0413	ctx-rh-entorhinal	0.0422	Left-Amygdala	0.0387
	ctx-rh-entorhinal	0.0340	Right-Amygdala	0.0339	Right-Hippocampus	0.0359
	Right-Amygdala	0.0302	ctx-rh-supramarginal	0.0261	ctx-lh-inferiortemporal	0.0306
	ctx-rh-middletemporal	0.0289	Left-Amygdala	0.0247	ctx-rh-middletemporal	0.0301
	Feature	Importance	Feature	Importance	Feature	Importance
Older age group Ages 77-84	Left-Hippocampus	0.0593	Left-Hippocampus	0.0411	Left-Hippocampus	0.0932
	ctx-rh-entorhinal	0.0432	Left-Inf-Lat-Vent	0.0350	Left-Amygdala	0.0451
	ctx-lh-middletemporal	0.0357	ctx-rh-entorhinal	0.0291	Right-Amygdala	0.0417
	Left-Amygdala	0.0333	ctx-lh-inferiortemporal	0.0286	ctx-rh-entorhinal	0.0345
	ctx-lh-entorhinal	0.0330	ctx-lh-entorhinal	0.0281	ctx-lh-middletemporal	0.0311
	ctx-lh-inferiortemporal	0.0304	ctx-lh-supramarginal	0.0263	Right-Hippocampus	0.0304
Total execution time for Stratified K-fold: **1.34 minutes** (80.67 seconds)						

To evaluate the consistency of identifying these top contributors, we analyzed the overlap of features in the same subgroup across all three tables. Out of the 54 features examined, derived from the top 6 contributing features within each of the nine subgroups, 39 were consistently present across all three tables. This indicates a high consistency score of 72.22% among the features analyzed. This

Table 3.8: Comparison of Top Contributing Features using Leave-One-Out Cross-Validation

Groups	Unified Sex (M+F)		Males (M)		Females (F)	
	Feature	Importance	Feature	Importance	Feature	Importance
Unified age group Ages 69-84	Left-Hippocampus	0.0618	Left-Hippocampus	0.05797	Left-Hippocampus	0.0693
	Left-Amygdala	0.0390	ctx-rh-entorhinal	0.04058	Right-Hippocampus	0.0384
	Right-Hippocampus	0.0296	Right-Amygdala	0.03326	Left-Amygdala	0.0380
	ctx-rh-entorhinal	0.0287	Left-Amygdala	0.03161	ctx-lh-entorhinal	0.0329
	Right-Amygdala	0.0279	ctx-lh-inferiorparietal	0.03147	ctx-rh-entorhinal	0.0255
	ctx-lh-entorhinal	0.0270	ctx-lh-middletemporal	0.02730	ctx-lh-middletemporal	0.0247
	Feature	Importance	Feature	Importance	Feature	Importance
Younger age group Ages 69-76	Left-Hippocampus	0.0604	Right-Hippocampus	0.0482	Left-Hippocampus	0.0560
	Left-Amygdala	0.0370	Left-Hippocampus	0.0421	Left-Amygdala	0.0419
	Right-Hippocampus	0.0340	ctx-rh-entorhinal	0.0418	ctx-lh-middletemporal	0.0412
	ctx-rh-entorhinal	0.0292	Right-Amygdala	0.0361	ctx-lh-entorhinal	0.0331
	ctx-lh-entorhinal	0.0273	Left-Amygdala	0.0278	Right-Hippocampus	0.0306
	Right-Amygdala	0.0250	ctx-lh-middletemporal	0.0240	ctx-rh-middletemporal	0.0294
	Feature	Importance	Feature	Importance	Feature	Importance
Older age group Ages 77-84	Left-Hippocampus	0.0673	Left-Hippocampus	0.0385	Left-Hippocampus	0.0888
	ctx-rh-entorhinal	0.0449	Left-Inf-Lat-Vent	0.0337	Right-Amygdala	0.0435
	ctx-lh-entorhinal	0.0342	ctx-lh-inferiortemporal	0.0336	Left-Amygdala	0.0382
	ctx-lh-inferiortemporal	0.0331	ctx-rh-entorhinal	0.0330	ctx-rh-entorhinal	0.0364
	Left-Amygdala	0.0308	Left-Amygdala	0.0301	ctx-lh-middletemporal	0.0301
	ctx-lh-middletemporal	0.0294	ctx-lh-entorhinal	0.0246	ctx-lh-entorhinal	0.0300
Total execution time for Leave-One-Out: **23.73 minutes**						

consistency underscores the dataset's influence on determining the top contributing features rather than the specific validation techniques, enhancing confidence in the predictive models. Table 3.9 presents the 39 features that consistently emerged as the top contributors across all subgroups and different validation techniques. Refer to the Supplementary section in Chapter (5) for detailed information on the performance metrics and top contributing features obtained through various cross-validation techniques for all subgroups.

Analyzing Table 3.9, we observe consistent involvement of specific brain regions across different age groups and sex. For instance, the left hippocampus consistently decreases in volume in both males and females across all age groups. Additionally, younger males (aged 69-76) and older females (aged 77-84) exhibit substantial volume reductions in the right amygdala. The younger (69-76) age group shows substantial volume decreases in the right hippocampus, highlighting its importance in the early stages of AD across the sexes. Furthermore, older males (aged 77-84) exhibit notable volume reductions in the left inferior temporal cortex region. These findings emphasize variations in brain region predictors based on age and sex.

Table 3.9: Consistent Top Contributing Features using K-fold, Stratified K-fold, and Leave-One-Out Cross-Validation

Groups	Unified Sex (M+F)	Males (M)	Females (F)
Unified age group Ages 69-84	Left-Hippocampus Left-Amygdala Right-Amygdala ctx-rh-entorhinal Right-Hippocampus	Left-Hippocampus Left-Amygdala ctx-rh-entorhinal ctx-lh-inferiorparietal	Left-Hippocampus Left-Amygdala Right-Hippocampus ctx-lh-entorhinal
Younger age group Ages 69-76	Left-Hippocampus Right-Hippocampus Left-Amygdala ctx-rh-entorhinal	Right-Hippocampus Left-Hippocampus Right-Amygdala ctx-rh-entorhinal Left-Amygdala	Left-Hippocampus Left-Amygdala ctx-lh-middletemporal
Older age group Ages 77-84	Left-Hippocampus ctx-rh-entorhinal Left-Amygdala ctx-lh-entorhinal ctx-lh-inferiortemporal	Left-Hippocampus Left-Inf-Lat-Vent ctx-lh-inferiortemporal ctx-lh-entorhinal ctx-rh-entorhinal	Left-Hippocampus Left-Amygdala ctx-lh-middletemporal Right-Amygdala

3.4.4 Neuroanatomical Trends in AD

In studying Table 3.9, the neuroanatomical trends in AD indicate consistent volume reductions across all brain regions, except for the lateral ventricle, which enlarges. These trends highlight sex-specific, age-related, and regional influences on AD.

Sex-Specific Trends:

- **Similarities observed in both males and females:**
  - The left hippocampus and left amygdala are key predictors in the Unified age group (69-84).
- **Differences observed across:**
  - Unified age group (69-84):
    - The ctx-lh-entorhinal (left entorhinal cortex) is a prominent predictor in females, while the ctx-lh-inferiorparietal (left inferior parietal cortex) is a prominent predictor in males.

  - Younger age group (69-76):
    - The ctx-rh-entorhinal (right entorhinal cortex) is a prominent predictor in males.
  - Older age group (77-84):
    - The ctx-lh-inferiortemporal (left inferior temporal cortex) is a leading predictor in males, while the ctx-lh-middletemporal (left middle temporal cortex) exhibits a higher ranking in females.

Age-Related Trends:

- **Similarities Observed in both younger (69-76) and older (77-84) age groups:**
  - The left hippocampus, left amygdala, and right entorhinal cortex are prominent predictors in the Unified Sex group (M+F).
  - The right entorhinal cortex is a prominent predictor in males, while the left middle temporal cortex shows prominence in females.
- **Differences Observed across:**
  - Unified Sex group (M+F):
    - The left inferior temporal cortex is a key predictor in the older age group (77-84).
  - Males only:
    - The left entorhinal cortex is a prominent predictor in the older age group (77-84).
    - The right amygdala and right hippocampus are key predictors in the younger age group (69-76).

  - Females only:
    - The right amygdala is a prominent predictor in the older age group (77-84).

Distinctive Regional Trends:

- Across both sexes and age groups (69-76 and 77-84), the left hippocampus and left amygdala consistently show higher rankings, implying their substantial role as key contributors to AD.
- The left middle temporal cortex is a prominent predictor among females across both younger (69-76) and older (77-84) age groups, suggesting a potential female-specific influence.
- Influence of the right entorhinal cortex is observed predominantly in males across both younger (69-76) and older (77-84) age groups, suggesting a potential male-specific influence.
- The right amygdala and right hippocampus are prominent predictors in younger males (aged 69-76) but not in the older male group (aged 77-84). Conversely, the right amygdala exhibits the opposite trend in females, being relevant in the older female group (aged 77-84) but not in the younger female group (aged 69-76). This observation suggests a potential interaction between age and sex affecting the right amygdala region.
- Influence of the left inferior temporal cortex and Left-Inf-Lat-Vent is male-specific in the older age group (77-84).
- The left entorhinal cortex is a prominent predictor in females and is also observed in older males (77-84).

The left inferior lateral ventricle (LILV) is a cerebrospinal fluid-filled cavity crucial for maintaining brain homeostasis and providing cushioning for brain structures. Enlargement of the LILV can indicate brain atrophy or neurodegenerative diseases as surrounding brain tissue shrinks, creating more space within the ventricles. This enlargement across cognitive states is evident in our results in Figure 3.4.

In Figure 3.4, the LILV shows a marked increase in volume from CN to AD, indicating ventricular enlargement with disease progression. The cortical regions (left inferior parietal cortex, left inferior temporal cortex, and left middle temporal cortex) demonstrate a decreasing volume trend from CN to AD, highlighting cortical atrophy as cognitive impairment advances. This trend is similar to the consistent volume decrease observed in the left and right hippocampus, amygdala, and entorhinal cortex from CN to MCI and AD, as seen in Figure 3.3. The left inferior parietal cortex and left inferior temporal cortex show a moderate volume decrease from CN to MCI and a more substantial reduction from MCI to AD, while the left middle temporal cortex exhibits a consistent decrease across all groups. This cortical volume reduction indicates neurodegenerative processes affecting these areas, known for their roles in spatial awareness, visual perception, language comprehension, and memory integration.

3.4.5 Insights from Visual Representations

Comprehensive Overview

Occurrences denote how frequently certain brain regions are observed in the analyses. Figure 3.5 provides a holistic view of brain region occurrences across subgroups and validation techniques. The results presented in Tables 3.6, 3.7, and 3.8 are visually depicted in a bar plot (Figure 3.5). This plot provides a holistic view of brain region occurrences across subgroups and validation techniques.

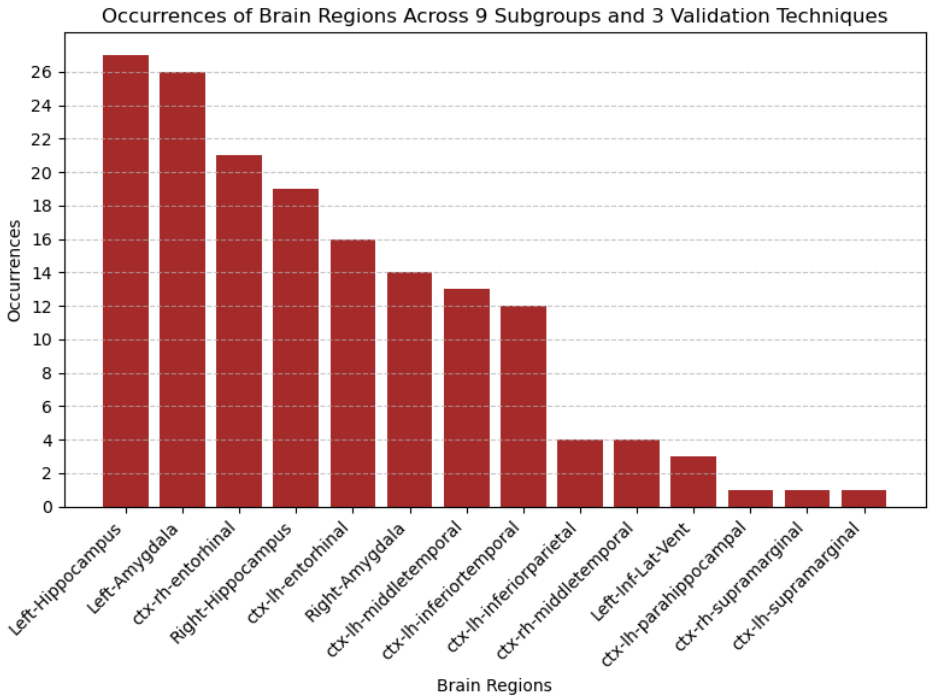
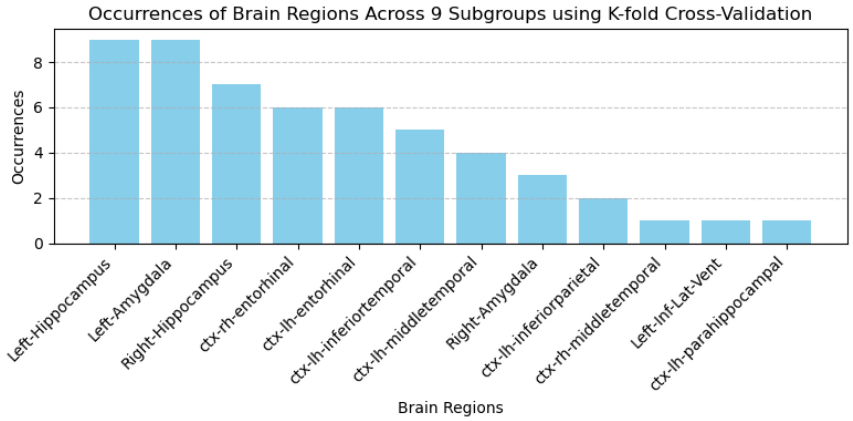


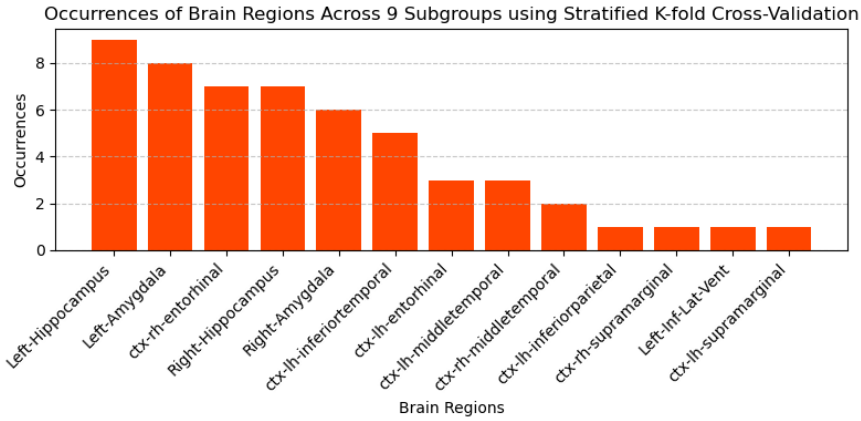
Figure 3.5: All occurrences of brain regions that drive AD across 9 subgroups using a combination of three validation techniques.

Detailed Occurrences plots

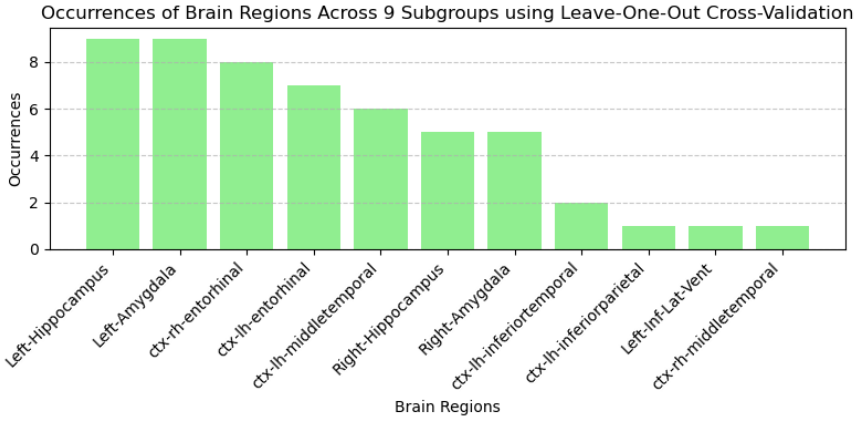
Figure 3.6 offers specific insights into the occurrence of brain regions across all subgroups and within individual validation techniques.



(a)



(b)



(c)

Figure 3.6: Most common high-ranking brain regions that contribute to AD found across 9 subgroups using different validation techniques (a) K-Fold (b) Stratified K-Fold (c) Leave-One-Out.

Freeview Visualization of Unique Brain Regions

Figure 3.7 utilizes FreeSurfer's Freeview tool to depict distinct brain regions identified in Figure 3.5 visually, offering insights into their precise anatomical locations within the brain.

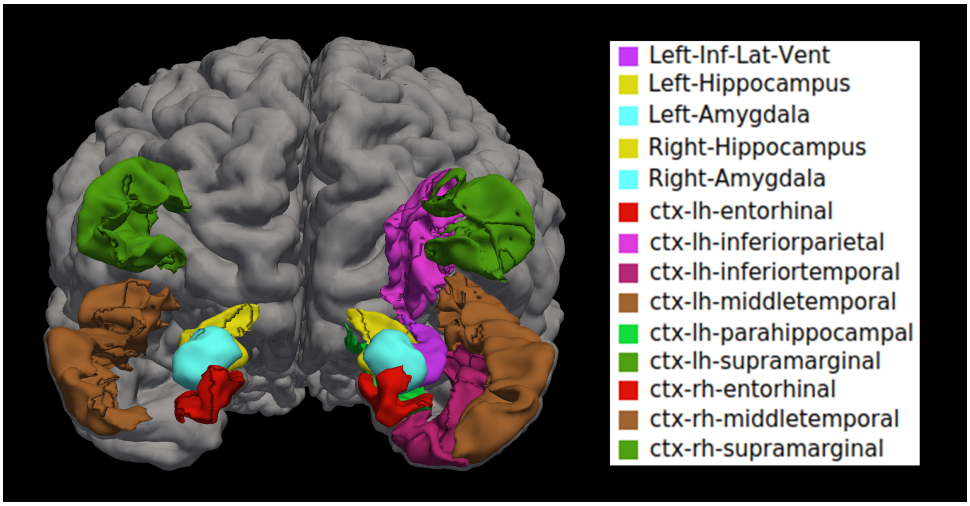


Figure 3.7: Visualization of unique brain regions that undergo substantial structural changes in Alzheimer's Disease across subgroups and validation techniques using FreeSurfer's Freeview.

Prominence of Highest Ranking Brain Regions

Focusing on the top six brain regions (Figure 3.8), we observe that the left hippocampus prominently features across all subgroups and validation techniques. Overall, the hippocampus, amygdala, and entorhinal cortex emerge as prominent predictors with decreased volume as Alzheimer's Disease progresses. These regions are visualized on Freesurfer in Figure 3.9.

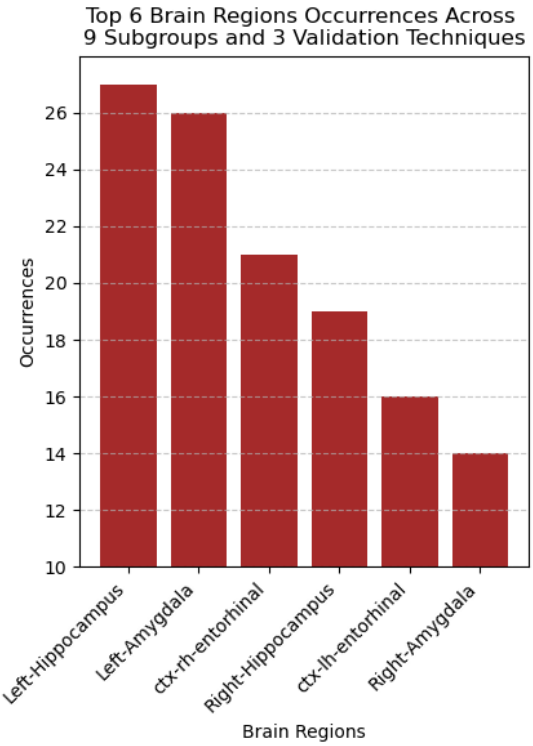


Figure 3.8: Top 6 occurrences of brain regions.

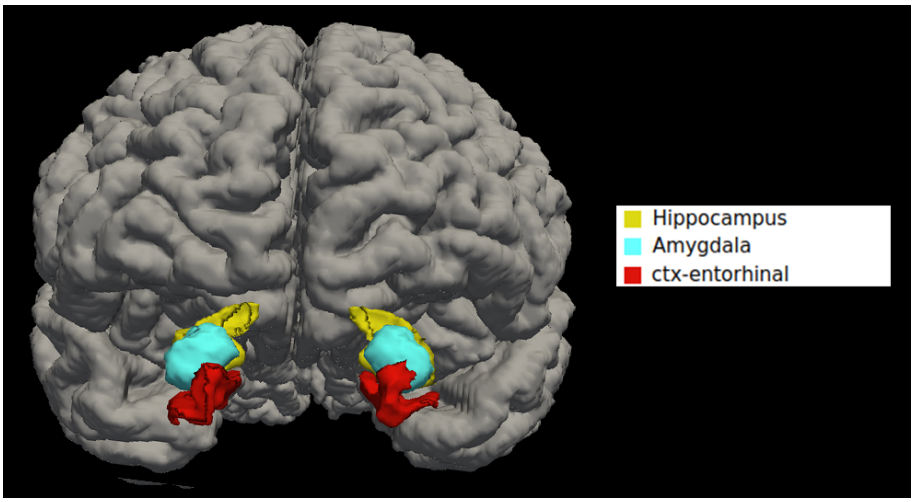


Figure 3.9: Visualization of top 6 brain regions - the hippocampus, amygdala, and entorhinal cortex on Freeview.

Chapter 4

Discussion

The primary objective of this project is to accurately predict CN, MCI, and AD diagnoses using neuroanatomical data from structural MRI scans in sex- and age-specific groups. Using the Random Forest (RF) algorithm and three cross-validation (CV) techniques, we achieved accuracies of 0.9086 for K-fold CV, 0.9187 for Stratified K-fold CV, and 0.9287 for LOOCV. Regressions were used to identify key brain regions undergoing structural changes in AD. Our analysis revealed that the hippocampus, amygdala, and entorhinal cortex consistently exhibited decreased volume in AD across various demographic groups. Sex-specific and age-specific patterns of brain region changes were observed, aiding in developing personalized diagnostic and therapeutic approaches. These insights can also inform tailored therapies or preventive strategies that consider the unique anatomical characteristics of different demographic groups. This ultimately enhances treatment effectiveness and improves patient outcomes.

4.1 Potential neuroanatomical markers of MCI and AD

We aimed to identify structurally affected brain regions in individuals aged 69-84 to capture the impact of the disease across this age range. Instead of focusing on disease progression with longitudinal scans using only CN and AD data, we included cross-sectional scans of CN, AD, and MCI patients to detect any structural changes. This approach allowed us to identify brain regions where changes appear or disappear from CN to MCI and from MCI to AD, providing a more comprehensive view. The six highest-ranked features, i.e., the left and right hippocampus, amygdala, and entorhinal cortex, represent the top three brain regions relevant to AD. We will now explore their roles in the disease.

4.1.1 Role of hippocampus and amygdala

Our analysis revealed consistent structural characteristics in the left hippocampus and left amygdala across both sexes and age groups (69-76 and 77-84), indicating their prominent role in AD. The right hippocampus and right amygdala were key predictors in younger males (69-76), but their relevance diminished in the older male group (77-84). Conversely, the right amygdala showed relevance in older females (77-84) but not in younger females (69-76), suggesting age and sex interactions in its involvement in AD pathogenesis.

The hippocampus

The hippocampus is essential for learning and memory, with its sub-regions contributing to the formation of episodic memory [97], [98]. In the initial stages of AD, rapid tissue loss occurs in the hippocampus, leading to functional disconnection

from other brain regions [99]. This damage disrupts the hippocampus's ability to function correctly, leading to memory loss and cognitive decline, which are hallmark symptoms of AD. Additionally, the hippocampus is involved in various cognitive functions beyond memory, such as emotional regulation and spatial orientation, and its dysfunction contributes to the broader cognitive impairment observed in AD.

The amygdala

The amygdala, integral to the limbic system, is linked to emotional disorders like autism, anxiety disorders, and AD, particularly in the context of fear. In the early stages of AD, both the amygdala and hippocampus change, leading to personality changes and emotional abnormalities like anxiety, mania, and irritability [100]. Because of early amygdala damage, mild stages of AD often exhibit neuropsychiatric symptoms. Eventually, about 80% of AD patients experience symptoms like hallucinations, delusions, paranoia, anxiety, agitation, and mood disturbances as the disease progresses [101], [102].

4.1.2 Role of entorhinal cortex

Our analysis revealed compelling trends in the entorhinal cortex concerning sex and age. While the right entorhinal cortex consistently showed volume reduction across both younger (age 69-76) and older (age 77-84) males, it did not exhibit this pattern in younger females (age 69-76). This suggests a male-specific influence in the right entorhinal cortex. In contrast, the left entorhinal cortex showed volume reduction in females across the unified age group (69-84). This trend was only observed for males in the older age group (77-84), indicating a potential age-specific pattern. These findings suggest complex interactions between sex and age in the entorhinal cortex's involvement in AD.

The entorhinal cortex serves as an initial focal point for the deposition of plaques and tangles in AD [103]. Plaques are abnormal clusters of beta-amyloid protein fragments that accumulate between nerve cells in the brain, while tau tangles are twisted tau protein fibers that form inside the brain cells. AD is characterized by these distinctive brain abnormalities — amyloid-β plaques and tau protein neurofibrillary tangles — which actively influence the neurodegenerative process [104]. During the early stages of AD, there is a buildup of tau protein within the entorhinal cortex, which propagates to the hippocampus [105]. Autopsy-based anatomical and histological investigations of Alzheimer's-afflicted brains have shown a progressive pattern of neurodegeneration, commencing in the second layer of the entorhinal cortex and gradually advancing to encompass the hippocampus, temporal cortex, frontoparietal cortex, and subcortical nuclei. The entorhinal cortex in individuals with AD experiences impaired function in processing sensory information and transmitting it for memory consolidation [106], [107]. This impairment results in substantial memory decline and difficulties with spatial navigation.

Our results contribute to the existing literature by providing detailed insights into the neurodegenerative patterns of AD. The observed volume reductions in the hippocampus, amygdala, and entorhinal cortex highlight their significance in AD, consistent with previous studies identifying these areas as critical for memory and emotional regulation. The sex- and age-specific trends in our findings demonstrate how AD affects different demographics, revealing variations in neurodegeneration based on sex and age. By identifying these key predictors, our study enhances the understanding of AD pathology.

4.2 Sex and Age Subgroup Analyses

Sex-based differences are prominent in the prevalence and progression of AD symptoms. In 2024, almost two-thirds of Americans with Alzheimer's are women; of the 6.9 million people aged 65 and older with Alzheimer's dementia, 4.2 million are women, and 2.7 million are men [23], [24]. Additionally, the lifetime risk for AD at age 45 is approximately 20% for women and 10% for men [23]. Males exhibit slower structural loss but faster age-related brain volume decline than females, a pattern that aligns with the distinct regional influences identified in AD across sexes and age groups. For instance, the male-specific prominence of the right entorhinal and the female-specific influence of the left middle temporal suggest potential neuroanatomical factors contributing to these sex-based differences.

Men, with their larger head size and cerebral brain volume compared to women, approximately 10% larger [108], may be less vulnerable to AD-related pathological factors, experiencing slower structural loss [109]. Yet, they also tend to undergo faster age-related brain volume decline compared to women [110]. Women are more likely to receive a clinical diagnosis of AD at similar levels of pathology than men [111]. Upon AD diagnosis, men often display reduced atrophy in various brain regions [112]. Women generally show a higher proportion of gray matter in various brain regions, whereas men tend to have a higher percentage of white matter [113]. These differences are often attributed to sex chromosomes and sex hormones, although the precise mechanism by which sex hormones affect brain structures remains unclear [109], [114].

As of 2024, in the United States, the percentage of people with Alzheimer's dementia increases with age: 5.0% of people aged 65 to 74, 13.2% of people aged 75 to 84, and 33.4% of people aged 85 and older are affected [23], [24]. In fact, data reveals

that annually, 4 out of 1,000 individuals aged 65 to 74, 32 out of 1,000 aged 75 to 84, and 76 out of 1,000 aged 85 and older develop AD [115]. Additionally, an estimated 8% to 11% of Americans aged 65 and older, totaling approximately 5 to 7 million individuals, may experience MCI due to AD [116]. Thus, the chance of developing AD increases with age. This aligns with our study's identification of age-specific neuroanatomical contributors, such as the male-specific involvement of the left entorhinal in the older age group (77-84). Most AD cases are diagnosed after the age of 65, with the prevalence becoming particularly concerning beyond this point, as the number of affected individuals doubles every five years. Among those aged 65-84, roughly one in thirteen has AD, while the number jumps to one in three for those over 85 [117]. Early-onset AD, affecting individuals under 65, is much less frequent, representing less than 10% of all cases [118]. The presence of the apolipoprotein E (APOE-ε4) allele, a variant of the apolipoprotein E gene, substantially increases the risk of AD in women compared to men [119], [120]. Women in their sixties who carry one or two copies of this gene variant are more likely to develop AD than their male counterparts [121].

4.3 Medications for AD

AD is incurable, but medications can provide relief by managing symptoms or potentially slowing disease progression. AD progressively damages and destroys nerve cells in the brain, leading to memory loss, reasoning impairment, and other cognitive declines. The objective is to decelerate this cognitive decline, enhancing the quality of life for individuals with Alzheimer's. Here's an overview of the two primary categories of drugs [122]:

4.3.1 Symptomatic Medications

These medications are designed to alleviate symptoms such as memory loss and confusion. However, it's important to note that these medicines are not approved or recommended for treating MCI.

1. **Cholinesterase Inhibitors:** These drugs prevent the breakdown of acetylcholine, a vital neurotransmitter involved in memory and learning.

   - **Donepezil (Aricept®):** Approved for all stages of AD.
   - **Rivastigmine (Exelon®):** Effective in managing mild to moderate AD symptoms.
   - **Galantamine (Razadyne®):** Approved for use in mild to moderate AD cases.

   While generally well-tolerated, these medications may lead to side effects such as nausea, diarrhea, loss of bladder control, muscle cramps, muscle twitching, and weight loss. Taking the medication at night may also result in vivid dreams [123].

2. **Glutamate Regulators:** These medications regulate glutamate, another neurotransmitter, to protect nerve cells from excessive activity that can harm them in AD patients.

   - **Memantine (Namenda®):** The only FDA-approved drug in this category for moderate to severe AD. It may slow cognitive decline and offer some nerve cell protection. Like cholinesterase inhibitors, Memantine is not a cure; side effects like dizziness, headache, confusion, hallucinations, agitation, and constipation can occur.

Our analysis reveals a strong association between AD and specific brain regions involved in memory, emotion, and spatial orientation. The effectiveness of these symptomatic medications may vary depending on which areas are affected in each patient. For example, cholinesterase inhibitors that boost acetylcholine levels, which affect cortical memory function [124], might benefit patients with substantial hippocampal involvement, given the hippocampus's role in memory [125], [126]. Similarly, glutamate regulators like Memantine could be more beneficial for patients with damage in the entorhinal cortex [127], where excessive neural activity can worsen cognitive decline. Understanding these specific brain region predictors can help healthcare professionals tailor treatments to maximize benefits and manage side effects better.

4.3.2 Disease-Modifying Drugs (Under Investigation)

These newer drugs aim to modify the underlying disease process of AD. They can potentially slow disease progression:

1. **Anti-amyloid Treatments:** These medications aim to reduce the accumulation of beta-amyloid plaques, a characteristic protein abnormality in AD.

    - **Lecanemab (Leqembi®):** This drug is an example of anti-amyloid treatment, working to remove beta-amyloid from the brain. It is approved by the FDA as an intravenous infusion for individuals in the early stages of AD with confirmed elevated brain amyloid levels. Beta-amyloid accumulates early in the entorhinal cortex and hippocampus, leading to cognitive deficits and synaptic dysfunction [128]. Therefore, Lecanemab targets these critical regions. However, Lecanemab is not a cure, and its long-term effects and safety are still being investigated.

Safety concerns were prominent in clinical trials [129], with 12.6% of participants developing brain edema and 17.3% experiencing brain hemorrhage. Some adverse events were severe enough to lead to discontinuation of the trial, and there were reports of deaths potentially associated with the drug, particularly when taken alongside anticoagulants or thrombolysis [130]. In the BAN2401-G000-201 trial [131], a 10-mg/kg biweekly dose of Lecanemab showed a 64% probability of outperforming placebo by 25% on the Alzheimer's Disease Composite Score (ADCOMS) at 12 months, which did not meet the 80% threshold for the primary endpoint. However, at 18 months, Lecanemab demonstrated substantial reductions in brain amyloid and clinical decline across multiple endpoints, like ADCOMS, ADAS-Cog14, and CDR-SB, with supportive changes in cerebrospinal fluid biomarkers and a 9.9% incidence of amyloid-related imaging abnormalities (ARIA). A systematic review and meta-analysis [132] involving 3,108 AD patients further supported these findings, showing statistically significant benefits in stabilizing or slowing cognitive decline as measured by CDR-SB, ADCOMS, and ADAS-cog scores, although the treatment increased the risk of ARIA.

Potential side effects of taking Lecanemab may include brain swelling or bleeding, and tests such as PET scans or spinal taps are required before initiating treatment [133]. Due to the risk of brain swelling or bleeding (known as ARIA), patients receiving Lecanemab require regular MRI scans for monitoring. These scans are recommended before the fifth, seventh, and fourteenth infusions and after one year of treatment. Rarely, these adverse effects may lead to symptoms such as headache, confusion, dizziness, changes in vision, nausea, and difficulty walking.

- **Kisunla (donanemab-azbt):** On July 2, 2024, the FDA approved Kisunla, another anti-amyloid treatment for AD. It is administered as an intravenous infusion every four weeks and is intended for patients with MCI or mild dementia. Clinical trials demonstrated that Kisunla significantly reduced clinical decline on various AD scales compared to placebo, including the Integrated Alzheimer's Disease Rating Scale (iADRS) and the Clinical Dementia Rating Scale – Sum of Boxes (CDR-SB). However, there are risks associated with Kisunla, such as ARIA and infusion-related reactions. The FDA granted Kisunla Fast Track, Priority Review, and Breakthrough Therapy designations [134].

2. **Other Potential Drugs:** Ongoing research explores medications targeting various aspects of AD pathology, including tau protein accumulation and neuroinflammation.

Aducanumab, once approved for the treatment of early-stage AD, was later withdrawn from the market due to concerns about its effectiveness [135]. Similarly, Lecanemab, which the FDA approved, has not shown substantial efficacy in halting the progression of the disease or reversing any associated damage [136].

Transcranial magnetic stimulation (TMS) and transcranial direct current stimulation (tDCS) are emerging as promising non-pharmacological interventions for AD. While previous clinical trials have primarily focused on evaluating their effects on global cognition, the authors in [137] explore the specific cognitive function of memory, which AD profoundly impacts. Using multilevel random effect models, the authors assessed the efficacy and safety of both TMS and tDCS in memory deficits among AD patients. The findings indicate that depending on the targeted brain regions, both interventions positively affected memory symptoms in AD patients.

Specifically, rTMS over frontal regions and tDCS over temporal regions demonstrated improvements in memory abilities. rTMS exhibited a statistical tendency towards enhancing memory in the long term, while the prolonged effects of tDCS could not be fully assessed due to limited data. The analysis also highlighted the safety of both techniques in the AD population, with minimal reports of serious adverse events. This review suggested the potential of TMS and tDCS as viable therapeutic options for addressing memory deficits in AD patients.

Important Considerations: Our research highlights the varied neuroanatomical patterns in AD across different sex and age groups. The effectiveness of medications in slowing cognitive decline may differ among individuals based on their specific brain characteristics. For instance, regions like the hippocampus, amygdala, and entorhinal cortex are top contributors to AD and may respond uniquely to treatments. Recognizing these differences can help doctors tailor treatments to each person, maximizing benefits and reducing side effects. Therefore, seeking personalized advice from healthcare professionals is crucial to managing AD symptoms effectively. Maintaining a healthy lifestyle, including regular exercise, a balanced diet, cognitive stimulation, and good sleep hygiene, can substantially slow AD progression.

4.4 Challenges Faced

4.4.1 Dataset to BIDS Hierarchy Transformation

One major challenge we faced involved transforming the dataset into a BIDS hierarchy. We tested seven conversion tools and sought assistance from our lab's alumni, Justin and Ryan, who were experienced in this research area. Unfortunately, their guidance proved insufficient, and our breakthrough came from mastering the HeudiConv tool, which proved instrumental.

4.4.2 Building the Machine Learning Dataset

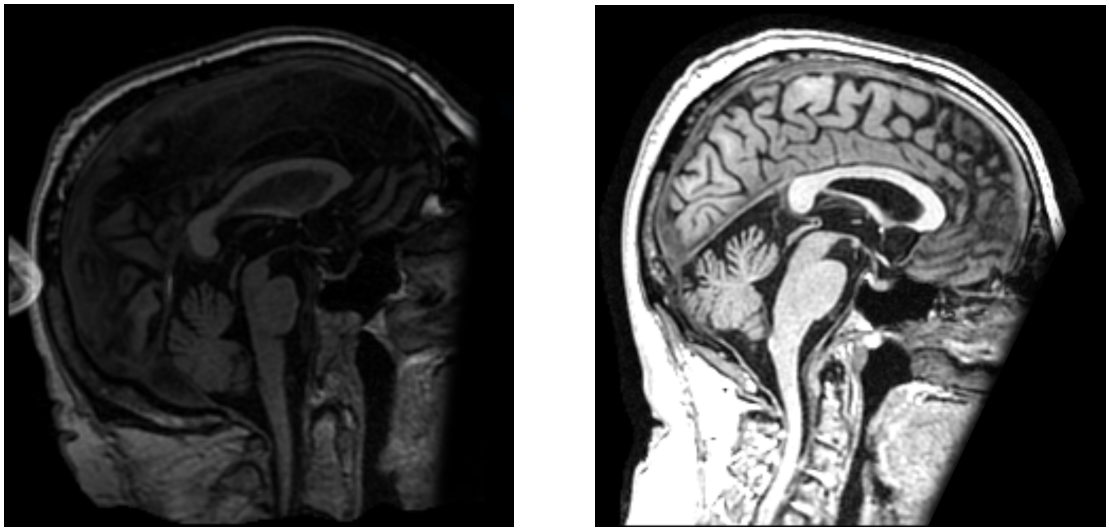
We encountered several challenges while setting up and using FastSurfer for our study:

1. **Missing Dependencies:** Initially, we faced errors related to missing modules like yacs and numpy. This was resolved by reinstalling FastSurfer within a dedicated Miniconda environment, ensuring all necessary dependencies were present.

2. **Cerebellum Parcellation Error:** We encountered an internal software error related to cerebellar parcellation. As this functionality wasn't critical for our research goals, we excluded the cerebellum from analysis using the `--no_cereb` flag in FastSurfer.

3. **IndexError:** While iteratively analyzing MRI scans from multiple subjects, we encountered an error (IndexError) for a specific scan. Upon inspection, this scan exhibited poor contrast, darkness, and missing brain regions (e.g., tissues), as seen in Figure 4.1. Due to these issues, FastSurfer could not process the scan. We addressed this by altogether removing the problematic scan from our dataset.

4. **Data Constraints:** Current anatomical findings are limited due to a lack of data on gray and white matter and functional data. This limited availability of data posed challenges to our research. Expanding our dataset could have further empowered our ML algorithm, potentially enhancing its accuracy in identifying AD neuroanatomical markers.



(a) (b)

Figure 4.1: Comparison of a discarded scan with a high-quality scan. (a) Defaced scan with missing brain regions. (b) High-quality scan.

4.5 Future Work

Building upon our research, future work can be directed toward several promising paths:

- Exploring atlases for Voxel-Based Morphometry [138],[139] and FSLeyes software [140] for brain volumetric extraction. Our initial experiments with FSLeyes yielded valuable insights in hippocampal volume extraction [141], [142] (see Figure 4.2). However, compared with FreeSurfer, we found the latter to be a more powerful tool suitable for our project's requirements.

- Exploring the link between anxiety and AD. We considered studying the link between anxiety and AD, but the literature review revealed no available anxiety-Alzheimer's comorbid dataset due to data security measures in hospitals, limiting our access. Establishing early intervention strategies targeting anxiety could potentially delay the onset of AD. However, limited access to compre-

hensive comorbid datasets may restrict statistically significant conclusions.

- Exploring alternative hyperparameter optimization techniques like Randomized Search and BayesSearchCV to enhance the tool's accuracy potentially. This could enhance model performance by identifying optimal parameters more efficiently than grid search, though it might increase computational cost and complexity.

- Studying the use of Adasyn or Recursive Feature Elimination (RFE) to streamline models by eliminating unnecessary features, offering an alternative to SMOTE. This could lead to simpler, faster models with improved generalization, though there is a risk of inadvertently removing important variables, potentially reducing model accuracy.

- Integrating deep learning methods with the Random Forest classifier may lead to a more precise tool for identifying the brain regions influencing AD. This could substantially enhance the precision of the tool, though increased model complexity could make interpretation harder and require more computational resources.

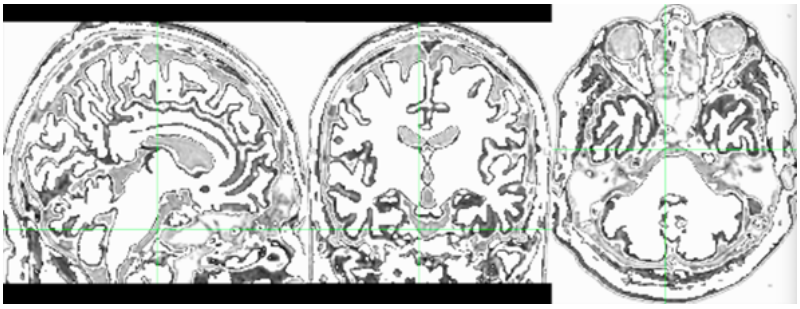
- Developing a clinical tool capable of analyzing a patient's MRI scan to identify specific regions in their brain associated with AD progression. This tool could aid clinicians in diagnosing AD earlier and more accurately, but navigating clinical validation and regulatory approval processes may be lengthy and challenging.

- Conducting a comprehensive comparison between Math and ML algorithms to determine a better AD predictor. This could improve the robustness and reliability of AD diagnosis tools, though the comparison study might be resource-

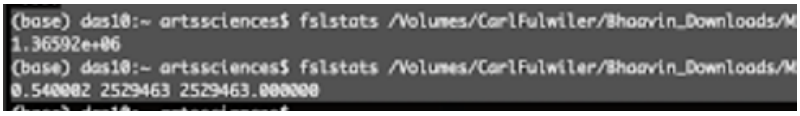
intensive and require extensive data for validation.

- Exploring the integration of fNIRS alongside MRI for enhanced early detection of MCI and AD, offering a cost-effective and comprehensive picture of brain health. This multimodal approach can track real-time brain activity during cognitive tasks and validate findings with structural data from MRI. Longitudinal studies using these tools will help identify early markers of cognitive decline and inform personalized intervention strategies. However, challenges such as standardizing fNIRS protocols across diverse populations and optimizing data integration with MRI remain critical for broader clinical applicability.

- Considering publication of promising results by preparing and submitting a manuscript to a prestigious conference for broader dissemination. This could enhance the visibility of our research, fostering collaboration and further advancements in the field.



(a)



(b)

Figure 4.2: Segmentation of brain tissues using FSLeyes. The white matter, gray matter, and cerebrospinal fluid are delineated. The volume of gray matter obtained in the brain was 2,529,463 mm^3. The atlas used here is the Harvard-Oxford cortical and subcortical structural atlas [143].

Other Subgroups in Alzheimer's Research

In addition to sex and age, other subgroups that can be explored to gain insights into the study of AD include:

1. **Genetic Variants:** Investigating the impact of specific genetic mutations or polymorphisms on AD risk and progression.

2. **Ethnicity and Race:** Examining differences in disease prevalence and outcomes among diverse racial and ethnic populations.

3. **Socioeconomic Status:** Exploring the influence of socioeconomic factors such as education, income, and access to healthcare on AD incidence and severity.

4. **Comorbidity Profiles:** Studying how other medical conditions or comorbidities may interact with AD pathology and affect disease outcomes.

In summary, images from the ADNI database were processed using Clinica, FreeSurfer, and FastSurfer software. The Random Forest Algorithm with K-fold, Stratified K-fold, and Leave-One-Out cross-validation techniques was employed to predict AD, MCI, and CN, achieving an average accuracy of 92.87% in detecting AD. Subgroup analyses were conducted to explore age and sex differences associated with AD. We learned that the hippocampus, amygdala, and entorhinal cortex are prominent predictors of Alzheimer's across both younger (69-76) and older (77-84) age groups. Additionally, specific brain regions demonstrate distinct volume decreases in males and females, such as the left middle temporal cortex in females and the right entorhinal cortex in males. We also reviewed symptomatic and disease-modifying drugs for AD. These findings highlight the critical role of specific brain regions in AD and underscore the potential for early detection and targeted interventions to address neurodegenerative changes.

Chapter 5

Supplementary Material

S1 - List of Brain Regions

95 brain regions of FastSurfer's 'segmentation-only' pipeline were used for training the AD classifier.

Cortical volumes: ctx-lh-caudalanteriorcingulate, ctx-lh-caudalmiddlefrontal, ctx-lh-cuneus, ctx-lh-entorhinal, ctx-lh-fusiform, ctx-lh-inferiorparietal, ctx-lh-inferiortemporal, ctx-lh-isthmuscingulate, ctx-lh-lateraloccipital, ctx-lh-lateralorbitofrontal, ctx-lh-lingual, ctx-lh-medialorbitofrontal, ctx-lh-middletemporal, ctx-lh-parahippocampal, ctx-lh-paracentral, ctx-lh-parsopercularis, ctx-lh-parsorbitalis, ctx-lh-parstriangularis, ctx-lh-pericalcarine, ctx-lh-postcentral, ctx-lh-posteriorcingulate, ctx-lh-precentral, ctx-lh-precuneus, ctx-lh-rostralanteriorcingulate, ctx-lh-rostralmiddlefrontal, ctx-lh-superiorfrontal, ctx-lh-superiorparietal, ctx-lh-superiortemporal, ctx-lh-supramarginal, ctx-lh-transversetemporal, ctx-lh-insula, ctx-rh-caudalanteriorcingulate, ctx-rh-caudalmiddlefrontal, ctx-rh-cuneus, ctx-rh-entorhinal, ctx-rh-fusiform, ctx-rh-inferiorparietal, ctx-rh-inferiortemporal, ctx-rh-isthmuscingulate, ctx-rh-lateraloccipital, ctx-rh-lateralorbitofrontal, ctx-rh-lingual, ctx-rh-

medialorbitofrontal, ctx-rh-middletemporal, ctx-rh-parahippocampal, ctx-rh-paracentral, ctx-rh-parsopercularis, ctx-rh-parsorbitalis, ctx-rh-parstriangularis, ctx-rh-pericalcarine, ctx-rh-postcentral, ctx-rh-posteriorcingulate, ctx-rh-precentral, ctx-rh-precuneus, ctx-rh-rostralanteriorcingulate, ctx-rh-rostralmiddlefrontal, ctx-rh-superiorfrontal, ctx-rh-superiorparietal, ctx-rh-superiortemporal, ctx-rh-supramarginal, ctx-rh-transversetemporal, ctx-rh-insula.

Subcortical gray matter volumes: Left-Cerebellum-Cortex, Left-Thalamus, Left-Caudate, Left-Putamen, Left-Pallidum, Brain-Stem, Left-Hippocampus, Left-Amygdala, Left-Accumbens-area, Left-VentralDC, Right-Cerebellum-Cortex, Right-Thalamus, Right-Caudate, Right-Putamen, Right-Pallidum, Right-Hippocampus, Right-Amygdala, Right-Accumbens-area, Right-VentralDC.

Ventricular system volumes: Left-Lateral-Ventricle, Left-Inf-Lat-Vent, 3rd-Ventricle, 4th-Ventricle, CSF, Left-choroid-plexus, Right-Lateral-Ventricle, Right-Inf-Lat-Vent, Right-choroid-plexus.

Subcortical white matter volumes: Left-Cerebral-White-Matter, Left-Cerebellum-White-Matter, Right-Cerebral-White-Matter, Right-Cerebellum-White-Matter, WM-hypointensities.

S2 - Performance metrics and Top Contributing Features of nine subgroups

Tables 3.6, 5.2, and 5.3 display the performance metrics and top contributing features obtained through K-fold, stratified K-fold, and leave-one-out cross-validation techniques.

Table 5.1: Comparison of Performance Metrics and Top Contributing Features using K-fold Cross-Validation

Age Group	Both sexes (M+F)				Males (M)				Females (F)			
	Aggregated Performance Metrics		**Top 6 Contributing Features**		**Aggregated Performance Metrics**		**Top 6 Contributing Features**		**Aggregated Performance Metrics**		**Top 6 Contributing Features**	
	Metric	Value	Feature	Importance	Metric	Value	Feature	Importance	Metric	Value	Feature	Importance
Ages 69 to 84	Accuracy	0.9086	Left-Hippocampus	0.0592	Accuracy	0.9415	Left-Hippocampus	0.04999	Accuracy	0.9357	Left-Hippocampus	0.0694
	Precision	0.9084	Left-Amygdala	0.0406	Precision	0.9414	Left-Amygdala	0.0381	Precision	0.9358	Left-Amygdala	0.0470
	Recall	0.9086	Right-Amygdala	0.0341	Recall	0.9415	ctx-rh-entorhinal	0.0355	Recall	0.9357	Right-Hippocampus	0.0352
	F1-score	0.9080	ctx-rh-entorhinal	0.0299	F1-score	0.9413	ctx-lh-inferiorparietal	0.0333	F1-score	0.9355	ctx-lh-entorhinal	0.0334
	Training accuracy	1.0	Right-Hippocampus	0.0296	Training accuracy	1	Right-Hippocampus	0.0323	Training accuracy	1.0	ctx-lh-inferiortemporal	0.0266
			ctx-lh-inferiortemporal	0.0251			ctx-lh-middletemporal	0.0289			ctx-lh-middletemporal	0.0261
	Aggregated Performance Metrics		**Top 6 Contributing Features**		**Aggregated Performance Metrics**		**Top 6 Contributing Features**		**Aggregated Performance Metrics**		**Top 6 Contributing Features**	
	Metric	Value	Feature	Importance	Metric	Value	Feature	Importance	Metric	Value	Feature	Importance
Ages 69 to 76	Accuracy	0.9051	Left-Hippocampus	0.0565	Accuracy	0.9315	Right-Hippocampus	0.0603	Accuracy	0.9403	Left-Hippocampus	0.0674
	Precision	0.9052	Right-Hippocampus	0.0356	Precision	0.9313	Left-Hippocampus	0.0443	Precision	0.9403	Left-Amygdala	0.0434
	Recall	0.9051	Left-Amygdala	0.0339	Recall	0.9315	Right-Amygdala	0.0423	Recall	0.9403	ctx-lh-parahippocampal	0.0333
	F1-score	0.9041	ctx-lh-entorhinal	0.0298	F1-score	0.9312	ctx-rh-entorhinal	0.0379	F1-score	0.9403	ctx-lh-middletemporal	0.0319
	Training accuracy	1.0	ctx-rh-entorhinal	0.0277	Training accuracy	1.0	Left-Amygdala	0.0302	Training accuracy	1.0	ctx-lh-entorhinal	0.0315
			ctx-rh-middletemporal	0.0274			ctx-lh-inferiorparietal	0.0215			ctx-lh-inferiortemporal	0.0292
	Aggregated Performance Metrics		**Top 6 Contributing Features**		**Aggregated Performance Metrics**		**Top 6 Contributing Features**		**Aggregated Performance Metrics**		**Top 6 Contributing Features**	
	Metric	Value	Feature	Importance	Metric	Value	Feature	Importance	Metric	Value	Feature	Importance
Ages 77 to 84	Accuracy	0.9098	Left-Hippocampus	0.0585	Accuracy	0.9282	Left-Hippocampus	0.0458	Accuracy	0.9365	Left-Hippocampus	0.0796
	Precision	0.9095	ctx-rh-entorhinal	0.0395	Precision	0.9284	Left-Inf-Lat-Vent	0.0357	Precision	0.9366	Left-Amygdala	0.0409
	Recall	0.9098	Left-Amygdala	0.0378	Recall	0.9282	ctx-lh-inferiortemporal	0.0323	Recall	0.9365	ctx-lh-middletemporal	0.0398
	F1-score	0.9095	ctx-lh-entorhinal	0.0373	F1-score	0.9281	ctx-lh-entorhinal	0.0320	F1-score	0.9360	Right-Amygdala	0.0389
	Training accuracy	1.0	ctx-lh-inferiortemporal	0.0336	Training accuracy	1.0	ctx-rh-entorhinal	0.0316	Training accuracy	1.0	ctx-lh-entorhinal	0.0344
			Right-Hippocampus	0.0243			Left-Amygdala	0.0308			Right-Hippocampus	0.0311
Total execution time for K-fold: **1.31 minutes** (78.76 seconds)												

Table 5.2: Comparison of Performance Metrics and Top Contributing Features using Stratified K-fold Cross-Validation

Age Group	Both sexes (M+F)				Males (M)				Females (F)			
	Aggregated Performance Metrics		**Top 6 Contributing Features**		**Aggregated Performance Metrics**		**Top 6 Contributing Features**		**Aggregated Performance Metrics**		**Top 6 Contributing Features**	
	Metric	Value	Feature	Importance	Metric	Value	Feature	Importance	Metric	Value	Feature	Importance
Ages 69 to 84	Accuracy	0.9187	Left-Hippocampus	0.0629	Accuracy	0.9503	Left-Hippocampus	0.0618	Accuracy	0.9357	Left-Hippocampus	0.0708
	Precision	0.9190	Left-Amygdala	0.0374	Precision	0.9502	ctx-rh-entorhinal	0.0347	Precision	0.9359	ctx-lh-entorhinal	0.0375
	Recall	0.9187	ctx-rh-entorhinal	0.0300	Recall	0.9503	Left-Amygdala	0.0293	Recall	0.9357	Left-Amygdala	0.0373
	F1-score	0.9178	Right-Hippocampus	0.0277	F1-score	0.9502	Right-Amygdala	0.0286	F1-score	0.9355	Right-Hippocampus	0.0337
	Training accuracy	1.0	Right-Amygdala	0.0254	Training accuracy	1.0	ctx-lh-inferiorparietal	0.0272	Training accuracy	1.0	ctx-lh-inferiortemporal	0.0275
			ctx-lh-inferiortemporal	0.0239			Right-Hippocampus	0.0263			Right-Amygdala	0.0248
	Aggregated Performance Metrics		**Top 6 Contributing Features**		**Aggregated Performance Metrics**		**Top 6 Contributing Features**		**Aggregated Performance Metrics**		**Top 6 Contributing Features**	
	Metric	Value	Feature	Importance	Metric	Value	Feature	Importance	Metric	Value	Feature	Importance
Ages 69 to 76	Accuracy	0.8962	Left-Hippocampus	0.0565	Accuracy	0.9583	Right-Hippocampus	0.0488	Accuracy	0.9353	Left-Hippocampus	0.0580
	Precision	0.8962	Left-Amygdala	0.0454	Precision	0.9583	Left-Hippocampus	0.0434	Precision	0.9355	ctx-lh-middletemporal	0.0502
	Recall	0.8962	Right-Hippocampus	0.0413	Recall	0.9583	ctx-rh-entorhinal	0.0422	Recall	0.9353	Left-Amygdala	0.0387
	F1-score	0.8957	ctx-rh-entorhinal	0.0340	F1-score	0.9582	Right-Amygdala	0.0339	F1-score	0.9351	Right-Hippocampus	0.0359
	Training accuracy	1.0	Right-Amygdala	0.0302	Training accuracy	1.0	ctx-rh-supramarginal	0.0261	Training accuracy	1.0	ctx-lh-inferiortemporal	0.0306
			ctx-rh-middletemporal	0.0289			Left-Amygdala	0.0247			ctx-rh-middletemporal	0.0301
	Aggregated Performance Metrics		**Top 6 Contributing Features**		**Aggregated Performance Metrics**		**Top 6 Contributing Features**		**Aggregated Performance Metrics**		**Top 6 Contributing Features**	
	Metric	Value	Feature	Importance	Metric	Value	Feature	Importance	Metric	Value	Feature	Importance
Ages 77 to 84	Accuracy	0.9116	Left-Hippocampus	0.0593	Accuracy	0.9224	Left-Hippocampus	0.0411	Accuracy	0.9365	Left-Hippocampus	0.0932
	Precision	0.9123	ctx-rh-entorhinal	0.0432	Precision	0.9222	Left-Inf-Lat-Vent	0.0350	Precision	0.9363	Left-Amygdala	0.0451
	Recall	0.9116	ctx-lh-middletemporal	0.0357	Recall	0.9224	ctx-rh-entorhinal	0.0291	Recall	0.9365	Right-Amygdala	0.0417
	F1-score	0.9107	Left-Amygdala	0.0333	F1-score	0.9222	ctx-lh-inferiortemporal	0.0286	F1-score	0.9363	ctx-rh-entorhinal	0.0345
	Training accuracy	1.0	ctx-lh-entorhinal	0.0330	Training accuracy	1.0	ctx-lh-entorhinal	0.0281	Training accuracy	1.0	ctx-lh-middletemporal	0.0311
			ctx-lh-inferiortemporal	0.0304			ctx-lh-supramarginal	0.0263			Right-Hippocampus	0.0304
Total execution time for Stratified K-fold: **1.34 minutes** (80.67 seconds)												

Table 5.3: Comparison of Performance Metrics and Top Contributing Features using Leave-One-Out Cross-Validation

Age Group	Both sexes (M+F)				Males (M)				Females (F)			
	Aggregated Performance Metrics		**Top 6 Contributing Features**		**Aggregated Performance Metrics**		**Top 6 Contributing Features**		**Aggregated Performance Metrics**		**Top 6 Contributing Features**	
	Metric	Value	Feature	Importance	Metric	Value	Feature	Importance	Metric	Value	Feature	Importance
Ages 69 to 84	Accuracy	0.9287	Left-Hippocampus	0.0618	Accuracy	0.9547	Left-Hippocampus	0.05797	Accuracy	0.9435	Left-Hippocampus	0.0693
	Precision	0.9284	Left-Amygdala	0.0390	Precision	0.9547	ctx-rh-entorhinal	0.04058	Precision	0.9435	Right-Hippocampus	0.0384
	Recall	0.9287	Right-Hippocampus	0.0296	Recall	0.9547	Right-Amygdala	0.03326	Recall	0.9435	Left-Amygdala	0.0380
	F1-score	0.9284	ctx-rh-entorhinal	0.0287	F1-score	0.9546	Left-Amygdala	0.03161	F1-score	0.9434	ctx-lh-entorhinal	0.0329
	Training accuracy	1.0	Right-Amygdala	0.0279	Training accuracy	1.0	ctx-lh-inferiorparietal	0.03147	Training accuracy	1.0	ctx-rh-entorhinal	0.0255
			ctx-lh-entorhinal	0.0270			ctx-lh-middletemporal	0.02730			ctx-lh-middletemporal	0.0247
	Aggregated Performance Metrics		**Top 6 Contributing Features**		**Aggregated Performance Metrics**		**Top 6 Contributing Features**		**Aggregated Performance Metrics**		**Top 6 Contributing Features**	
	Metric	Value	Feature	Importance	Metric	Value	Feature	Importance	Metric	Value	Feature	Importance
Ages 69 to 76	Accuracy	0.9205	Left-Hippocampus	0.0604	Accuracy	0.9435	Right-Hippocampus	0.0482	Accuracy	0.9303	Left-Hippocampus	0.0560
	Precision	0.9212	Left-Amygdala	0.0370	Precision	0.9435	Left-Hippocampus	0.0421	Precision	0.9307	Left-Amygdala	0.0419
	Recall	0.9205	Right-Hippocampus	0.0340	Recall	0.9435	ctx-rh-entorhinal	0.0418	Recall	0.9303	ctx-lh-middletemporal	0.0412
	F1-score	0.9200	ctx-rh-entorhinal	0.0292	F1-score	0.9430	Right-Amygdala	0.0361	F1-score	0.9304	ctx-lh-entorhinal	0.0331
	Training accuracy	1.0	ctx-lh-entorhinal	0.0273	Training accuracy	1.0	Left-Amygdala	0.0278	Training accuracy	1.0	Right-Hippocampus	0.0306
			Right-Amygdala	0.0250			ctx-lh-middletemporal	0.0240			ctx-rh-middletemporal	0.0294
	Aggregated Performance Metrics		**Top 6 Contributing Features**		**Aggregated Performance Metrics**		**Top 6 Contributing Features**		**Aggregated Performance Metrics**		**Top 6 Contributing Features**	
	Metric	Value	Feature	Importance	Metric	Value	Feature	Importance	Metric	Value	Feature	Importance
Ages 77 to 84	Accuracy	0.9263	Left-Hippocampus	0.0673	Accuracy	0.9397	Left-Hippocampus	0.0385	Accuracy	0.9460	Left-Hippocampus	0.0888
	Precision	0.9272	ctx-rh-entorhinal	0.0449	Precision	0.9396	Left-Inf-Lat-Vent	0.0337	Precision	0.9477	Right-Amygdala	0.0435
	Recall	0.9263	ctx-lh-entorhinal	0.0342	Recall	0.9397	ctx-lh-inferiortemporal	0.0336	Recall	0.9460	Left-Amygdala	0.0382
	F1-score	0.9257	ctx-lh-inferiortemporal	0.0331	F1-score	0.9396	ctx-rh-entorhinal	0.0330	F1-score	0.9455	ctx-rh-entorhinal	0.0364
	Training accuracy	1.0	Left-Amygdala	0.0308	Training accuracy	1.0	Left-Amygdala	0.0301	Training accuracy	1.0	ctx-lh-middletemporal	0.0301
			ctx-lh-middletemporal	0.0294			ctx-lh-entorhinal	0.0246			ctx-lh-entorhinal	0.0300
Total execution time for Leave-One-Out: **23.73 minutes**												

Bibliography

[1] A. Association, "2023 alzheimer's disease facts and figures," *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, vol. 19, no. 4, pp. 1598–1695, 2023. DOI: 10.1002/alz.13016.

[2] X. Li, X. Feng, X. Sun, N. Hou, F. Han, and Y. Liu, "Global, regional, and national burden of alzheimer's disease and other dementias, 1990-2019," *Frontiers in aging neuroscience*, vol. 14, p. 937 486, 2022, ISSN: 1663-4365. DOI: 10.3389/fnagi.2022.937486. [Online]. Available: https://europepmc.org/articles/PMC9588915.

[3] J. Olazarán, C. Carnero-Pardo, J. Fortea, *et al.*, "Prevalence of treated patients with alzheimer's disease: Current trends and covid-19 impact," *Alzheimer's Research & Therapy*, vol. 15, no. 1, p. 130, 2023.

[4] M. S. Baek, H.-K. Kim, K. Han, *et al.*, "Annual trends in the incidence and prevalence of alzheimer's disease in south korea: A nationwide cohort study," *Frontiers in Neurology*, vol. 13, p. 883 549, 2022.

[5] K. A. Matthews, W. Xu, A. H. Gaglioti, *et al.*, "Racial and ethnic estimates of alzheimer's disease and related dementias in the united states (2015–2060) in adults aged 65 years," *Alzheimer's & Dementia*, vol. 15, no. 1, pp. 17–24, 2019.

[6] M. Khojaste-Sarakhsi, S. S. Haghighi, S. F. Ghomi, and E. Marchiori, "Deep learning for alzheimer's disease diagnosis: A survey," *Artificial intelligence in medicine*, vol. 130, p. 102 332, 2022.

[7] V. Laganà, F. Bruno, N. Altomari, *et al.*, "Neuropsychiatric or behavioral and psychological symptoms of dementia (bpsd): Focus on prevalence and natural history in alzheimer's disease and frontotemporal dementia," *Frontiers in neurology*, vol. 13, p. 832 199, 2022.

[8] P. Marešová, J. Dolejs, H. Mohelska, and L. K. Bryan, "Cost of treatment and care for people with alzheimer's disease: A meta-analysis," *Current Alzheimer Research*, vol. 16, no. 14, pp. 1245–1253, 2019.

[9] A. Association *et al.*, "2018 alzheimer's disease facts and figures," *Alzheimer's & Dementia*, vol. 14, no. 3, pp. 367–429, 2018.

[10] A. R. Roda, G. Serra-Mir, L. Montoliu-Gaya, L. Tiessler, and S. Villegas, "Amyloid-beta peptide and tau protein crosstalk in alzheimer's disease," *Neural regeneration research*, vol. 17, no. 8, pp. 1666–1674, 2022.

[11] J. Bok, J. Ha, B. J. Ahn, and Y. Jang, "Disease-modifying effects of non-invasive electroceuticals on β-amyloid plaques and tau tangles for alzheimer's disease," *International Journal of Molecular Sciences*, vol. 24, no. 1, p. 679, 2022.

[12] R. Ismail, P. Parbo, L. S. Madsen, *et al.*, "The relationships between neuroinflammation, beta-amyloid and tau deposition in alzheimer's disease: A longitudinal pet study," *Journal of neuroinflammation*, vol. 17, pp. 1–11, 2020.

[13] T. J. Betthauser, R. L. Koscik, E. M. Jonaitis, *et al.*, "Amyloid and tau imaging biomarkers explain cognitive decline from late middle-age," *Brain*, vol. 143, no. 1, pp. 320–335, 2020.

[14] *Grey matter*, `https://my.clevelandclinic.org/health/body/24831-grey-matter`, Accessed: June 8, 2024.

[15] *Gray matter vs. white matter in the brain*, `https://www.spinalcord.com/blog/gray-matter-vs-white-matter-in-the-brain`, Accessed: June 8, 2024.

[16] C. R. Jack Jr, M. A. Bernstein, N. C. Fox, *et al.*, "The alzheimer's disease neuroimaging initiative (adni): Mri methods," *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 27, no. 4, pp. 685–691, 2008.

[17] K.-H. Nenning and G. Langs, "Machine learning in neuroimaging: From research to clinical practice," *Die Radiologie*, vol. 62, no. Suppl 1, pp. 1–10, 2022.

[18] Y. Varatharajah, V. K. Ramanan, R. Iyer, and P. Vemuri, "Predicting short-term mci-to-ad progression using imaging, csf, genetic factors, cognitive resilience, and demographics," *Scientific reports*, vol. 9, no. 1, p. 2235, 2019.

[19] W. Lin, T. Tong, Q. Gao, *et al.*, "Convolutional neural networks-based mri image analysis for the alzheimer's disease prediction from mild cognitive impairment," *Frontiers in neuroscience*, vol. 12, p. 777, 2018.

[20] M. Ahmadzadeh, G. J. Christie, T. D. Cosco, *et al.*, "Neuroimaging and machine learning for studying the pathways from mild cognitive impairment to alzheimer's disease: A systematic review," *BMC neurology*, vol. 23, no. 1, p. 309, 2023.

[21] K. R. Laws, K. Irvine, and T. M. Gale, "Sex differences in cognitive impairment in alzheimer's disease," *World journal of psychiatry*, vol. 6, no. 1, p. 54, 2016.

[22] K. R. Laws, K. Irvine, and T. M. Gale, "Sex differences in alzheimer's disease," *Current opinion in psychiatry*, vol. 31, no. 2, pp. 133–139, 2018.

[23] "2024 alzheimer's disease facts and figures," *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, vol. 20, no. 5, pp. 3708–3821, 2024. DOI: 10.1002/alz.13809. [Online]. Available: https://doi.org/10.1002/alz.13809.

[24] K. B. Rajan, J. Weuve, L. L. Barnes, E. A. McAninch, R. S. Wilson, and D. A. Evans, "Population estimate of people with clinical alzheimer's disease and mild cognitive impairment in the united states (2020–2060)," *Alzheimer's & dementia*, vol. 17, no. 12, pp. 1966–1975, 2021.

[25] M. C. Staff, "Alzheimer's disease," *Mayo Clinic*, 2023, https://www.mayoclinic.org/diseases-conditions/alzheimers-disease/in-depth/alzheimers/art-20048356, accessed on June 17, 2024.

[26] P.-P. Liu, Y. Xie, X.-Y. Meng, and J.-S. Kang, "History and progress of hypotheses and clinical trials for alzheimer's disease," *Signal transduction and targeted therapy*, vol. 4, no. 1, p. 29, 2019.

[27] P. Davies and A. Maloney, "Selective loss of central cholinergic neurons in alzheimer's disease," *The Lancet*, vol. 308, no. 8000, p. 1403, 1976.

[28] P. T. Francis, A. M. Palmer, M. Snape, and G. K. Wilcock, "The cholinergic hypothesis of alzheimer's disease: A review of progress," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 66, no. 2, pp. 137–147, 1999.

[29] D. Fotiou, A. Kaltsatou, D. Tsiptsios, and M. Nakou, "Evaluation of the cholinergic hypothesis in alzheimer's disease with neuropsychological methods," *Aging clinical and experimental research*, vol. 27, pp. 727–733, 2015.

[30] T. H Ferreira-Vieira, I. M Guimaraes, F. R Silva, and F. M Ribeiro, "Alzheimer's disease: Targeting the cholinergic system," *Current neuropharmacology*, vol. 14, no. 1, pp. 101–115, 2016.

[31] L. Håkansson, "Mechanism of action of cholinesterase inhibitors in alzheimer's disease," *Acta Neurologica Scandinavica*, vol. 88, no. S149, pp. 7–9, 1993.

[32] P. Anand and B. Singh, "A review on cholinesterase inhibitors for alzheimer's disease," *Archives of pharmacal research*, vol. 36, pp. 375–399, 2013.

[33] O. Jordana, K. L. Lanctôt, G. Mazereeuw, N. Herrmann, *et al.*, "Cholinesterase inhibitor discontinuation in patients with alzheimer's disease: A meta-analysis of randomized controlled trials," *The Journal of Clinical Psychiatry*, vol. 76, no. 11, p. 2613, 2015.

[34] Prescrire Int, "Drugs for alzheimer's disease: Best avoided. no therapeutic advantage," *Prescrire International*, vol. 21, no. 128, p. 150, Jun. 2012.

[35] D. J. Selkoe, "The molecular pathology of alzheimer's disease," *Neuron*, vol. 6, no. 4, pp. 487–498, 1991.

[36] J. Hardy and D. Allsop, "Amyloid deposition as the central event in the aetiology of alzheimer's disease," *Trends in pharmacological sciences*, vol. 12, pp. 383–388, 1991.

[37] L. S. Honig, B. Vellas, M. Woodward, *et al.*, "Trial of solanezumab for mild dementia due to alzheimer's disease," *New England Journal of Medicine*, vol. 378, no. 4, pp. 321–330, 2018.

[38] S. Ostrowitzki, R. A. Lasser, E. Dorflinger, *et al.*, "A phase iii randomized trial of gantenerumab in prodromal alzheimer's disease," *Alzheimer's research & therapy*, vol. 9, pp. 1–15, 2017.

[39] B. Frost, R. L. Jacks, and M. I. Diamond, "Propagation of tau misfolding from the outside to the inside of a cell," *Journal of Biological Chemistry*, vol. 284, no. 19, pp. 12 845–12 852, 2009.

[40] H. Braak and E. Braak, "Evolution of the neuropathology of alzheimer's disease," *Acta Neurologica Scandinavica*, vol. 94, no. S165, pp. 3–12, 1996.

[41] E. Braak, K. Griffing, K. Arai, J. Bohl, H. Bratzke, and H. Braak, "Neuropathology of alzheimer's disease: What is new since a. alzheimer?" *European archives of psychiatry and clinical neuroscience*, vol. 249, S14–S22, 1999.

[42] G. K. Wilcock, S. Gauthier, G. B. Frisoni, *et al.*, "Potential of low dose leuco-methylthioninium bis (hydromethanesulphonate)(lmtm) monotherapy for treatment of mild alzheimer's disease: Cohort analysis as modified primary outcome in a phase iii clinical trial," *Journal of Alzheimer's disease*, vol. 61, no. 1, pp. 435–457, 2018.

[43] P. Novak, R. Schmidt, E. Kontsekova, *et al.*, "Fundamant: An interventional 72-week phase 1 follow-up study of aadvac1, an active immunotherapy against tau protein pathology in alzheimer's disease," *Alzheimer's research & therapy*, vol. 10, pp. 1–16, 2018.

[44] R. H. Swerdlow and S. M. Khan, "A "mitochondrial cascade hypothesis" for sporadic alzheimer's disease," *Medical hypotheses*, vol. 63, no. 1, pp. 8–20, 2004.

[45] M. P. Mattson, B. Cheng, D. Davis, K. Bryant, I. Lieberburg, and R. E. Rydel, "Beta-amyloid peptides destabilize calcium homeostasis and render human cortical neurons vulnerable to excitotoxicity," *Journal of Neuroscience*, vol. 12, no. 2, pp. 376–389, 1992.

[46] C. Iadecola, "Neurovascular regulation in the normal brain and in alzheimer's disease," *Nature Reviews Neuroscience*, vol. 5, no. 5, pp. 347–360, 2004.

[47] C. K. Kim, Y. R. Lee, L. Ong, M. Gold, A. Kalali, and J. Sarkar, "Alzheimer's disease: Key insights from two decades of clinical trial failures," *Journal of Alzheimer's Disease*, vol. 87, no. 1, pp. 83–100, 2022.

[48] J. Cummings, "Lessons learned from alzheimer disease: Clinical trials with negative outcomes," *Clinical and translational science*, vol. 11, no. 2, p. 147, 2018.

[49] J. Wang, V. Logovinsky, S. B. Hendrix, *et al.*, "Adcoms: A composite clinical outcome for prodromal alzheimer's disease trials," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 87, no. 9, pp. 993–999, 2016.

[50] S. Tucker, C. Möller, K. Tegerstedt, *et al.*, "The murine version of ban2401 (mab158) selectively reduces amyloid-β protofibrils in brain and cerebrospinal fluid of tg-arcswe mice," *Journal of Alzheimer's Disease*, vol. 43, no. 2, pp. 575–588, 2015.

[51] V. A. Canady, "Fda approves new treatment for alzheimer's disease," *Mental Health Weekly*, vol. 33, no. 3, pp. 6–7, 2023.

[52] C. J. Swanson, Y. Zhang, S. Dhadda, *et al.*, "A randomized, double-blind, phase 2b proof-of-concept clinical trial in early alzheimer's disease with lecanemab, an anti-aβ protofibril antibody," *Alzheimer's research & therapy*, vol. 13, pp. 1–14, 2021.

[53] U. Food and D. Administration, *Fda converts novel alzheimer's disease treatment to traditional approval*, Accessed: 2024-06-29, 2023. [Online]. Available: https://www.fda.gov/news-events/press-announcements/fda-converts-novel-alzheimers-disease-treatment-traditional-approval.

[54] M. C. Irizarry, J. R. Sims, S. L. Lowe, *et al.*, "O4-08-06: Safety, pharmacokinetics (pk), and florbetapir f-18 positron emission tomography (pet) after multiple dose administration of ly3002813, a β-amyloid plaque-specific antibody, in alzheimer's disease (ad)," *Alzheimer's & Dementia*, vol. 12, P352–P353, 2016.

[55] S. Shcherbinin, C. D. Evans, M. Lu, *et al.*, "Association of amyloid reduction after donanemab treatment with tau pathology and clinical outcomes: The trailblazer-alz randomized clinical trial," *JAMA neurology*, vol. 79, no. 10, pp. 1015–1024, 2022.

[56] J. R. Sims, J. A. Zimmer, C. D. Evans, *et al.*, "Donanemab in early symptomatic alzheimer disease: The trailblazer-alz 2 randomized clinical trial," *Jama*, vol. 330, no. 6, pp. 512–527, 2023.

[57] C. H. Van Dyck, C. J. Swanson, P. Aisen, *et al.*, "Lecanemab in early alzheimer's disease," *New England Journal of Medicine*, vol. 388, no. 1, pp. 9–21, 2023.

[58] M. A. Mintun, A. C. Lo, C. Duggan Evans, *et al.*, "Donanemab in early alzheimer's disease," *New England Journal of Medicine*, vol. 384, no. 18, pp. 1691–1704, 2021.

[59] L.-K. Huang, S.-P. Chao, and C.-J. Hu, "Clinical trials of new drugs for alzheimer disease," *Journal of biomedical science*, vol. 27, pp. 1–13, 2020.

[60] L.-K. Huang, Y.-C. Kuan, H.-W. Lin, and C.-J. Hu, "Clinical trials of new drugs for alzheimer disease: A 2020–2023 update," *Journal of Biomedical Science*, vol. 30, no. 1, p. 83, 2023.

[61] L. Bloch and C. M. Friedrich, "Developing a machine learning workflow to explain black-box models for alzheimer's disease classification.," in *HEALTHINF*, 2021, pp. 87–99.

[62] L. Sørensen, C. Igel, A. Pai, *et al.*, "Differential diagnosis of mild cognitive impairment and alzheimer's disease using structural mri cortical thickness, hippocampal shape, hippocampal texture, and volumetry," *NeuroImage: Clinical*, vol. 13, pp. 470–482, 2017.

[63] L. Bloch and C. M. Friedrich, "Comparison of automated volume extraction with freesurfer and fastsurfer for early alzheimer's disease detection with machine learning," in *2021 IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS)*, IEEE, 2021, pp. 113–118.

[64] A. Shukla, R. Tiwari, and S. Tiwari, "Alzheimer's disease detection from fused pet and mri modalities using an ensemble classifier," *Machine Learning and Knowledge Extraction*, vol. 5, no. 2, pp. 512–538, 2023.

[65] V. S. Rallabandi, K. Tulpule, M. Gattu, A. D. N. Initiative, *et al.*, "Automatic classification of cognitively normal, mild cognitive impairment and alzheimer's disease using structural mri analysis," *Informatics in Medicine Unlocked*, vol. 18, p. 100 305, 2020.

[66] K. R. Gray, P. Aljabar, R. A. Heckemann, A. Hammers, D. Rueckert, A. D. N. Initiative, *et al.*, "Random forest-based similarity measures for multi-modal classification of alzheimer's disease," *NeuroImage*, vol. 65, pp. 167–175, 2013.

[67] A. Lebedev, E. Westman, G. Van Westen, *et al.*, "Random forest ensembles for detection and prediction of alzheimer's disease with a good between-cohort robustness," *NeuroImage: Clinical*, vol. 6, pp. 115–125, 2014.

[68] M. Velazquez, Y. Lee, and A. D. N. Initiative, "Random forest model for feature-based alzheimer's disease conversion prediction from early mild cognitive impairment subjects," *Plos one*, vol. 16, no. 4, e0244773, 2021.

[69] M. Song, H. Jung, S. Lee, D. Kim, and M. Ahn, "Diagnostic classification and biomarker identification of alzheimer's disease with random forest algorithm," *Brain sciences*, vol. 11, no. 4, p. 453, 2021.

[70] I. Arevalo-Rodriguez, N. Smailagic, M. R. i Figuls, *et al.*, "Mini-mental state examination (mmse) for the detection of alzheimer's disease and other dementias in people with mild cognitive impairment (mci)," *Cochrane database of systematic reviews*, no. 3, 2015.

[71] J. C. Morris, "Clinical dementia rating: A reliable and valid diagnostic and staging measure for dementia of the alzheimer type," *International psychogeriatrics*, vol. 9, no. S1, pp. 173–176, 1997.

[72] A. Routier, N. Burgos, M. Díaz, *et al.*, "Clinica: An open-source software platform for reproducible clinical neuroscience studies," *Frontiers in Neuroinformatics*, vol. 15, p. 689 675, 2021.

[73] K. J. Gorgolewski, T. Auer, V. D. Calhoun, *et al.*, "The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments," *Scientific data*, vol. 3, no. 1, pp. 1–9, 2016.

[74] Y. Halchenko, M. Goncalves, P. Velasco, *et al.*, *Nipy/heudiconv: V1.0.0*, version v1.0.0, Sep. 2023. DOI: 10.5281/zenodo.8364586. [Online]. Available: https://doi.org/10.5281/zenodo.8364586.

[75] X. Li, P. S. Morgan, J. Ashburner, J. Smith, and C. Rorden, "The first step for neuroimaging data analysis: Dicom to nifti conversion," *Journal of neuroscience methods*, vol. 264, pp. 47–56, 2016.

[76] Stanford, *BIDS Tutorial Series Part 2A*, https://reproducibility.stanford.edu/bids-tutorial-series-part-2a/, Accessed: October 2023, 2018.

[77] M. P. Zwiers, S. Moia, and R. Oostenveld, "Bidscoin: A user-friendly application to convert source data to brain imaging data structure," *Frontiers in Neuroinformatics*, vol. 15, p. 65, 2022.

[78] NILAB-UvA, *Bidsify*, https://github.com/NILAB-UvA/bidsify, Accessed: October 2023, 2019.

[79] jmtyszka, *Bidskit QuickStart*, https://github.com/jmtyszka/bidskit/blob/master/docs/QuickStart.md, Accessed: October 2023, 2023.

[80] SIMEXP, *Data2Bids*, https://github.com/SIMEXP/Data2Bids, Accessed: October 2023, 2019.

[81] A. Boré, S. Guay, C. Bedetti, S. Meisler, and N. GuenTher, *Dcm2bids*, version 3.1.1, Oct. 2023. DOI: 10.5281/zenodo.8436509. [Online]. Available: https://doi.org/10.5281/zenodo.8436509.

[82] benoitberanger, *niix2bids*, https://github.com/benoitberanger/niix2bids, Accessed: October 2023, 2022.

[83] dangom, *dac2bids*, https://github.com/dangom/dac2bids, Accessed: October 2023, 2017.

[84] U. Bashir, J. Yap, B. Rasuli, *et al.*, *Diffusion-weighted imaging*, Reference article, Accessed on 16 Mar 2024, 2024. [Online]. Available: https://doi.org/10.53347/rID-16718.

[85] B. Fischl, "Freesurfer," *Neuroimage*, vol. 62, no. 2, pp. 774–781, 2012.

[86] L. Henschel, S. Conjeti, S. Estrada, K. Diers, B. Fischl, and M. Reuter, "Fastsurfer-a fast and accurate deep learning based neuroimaging pipeline," *NeuroImage*, vol. 219, p. 117 012, 2020.

[87] K. J. Liang and E. S. Carlson, "Resistance, vulnerability and resilience: A review of the cognitive cerebellum in aging and neurodegenerative diseases," *Neurobiology of Learning and Memory*, vol. 170, p. 106 981, 2020.

[88] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: Synthetic minority over-sampling technique," *Journal of artificial intelligence research*, vol. 16, pp. 321–357, 2002.

[89] H. He, Y. Bai, E. A. Garcia, and S. Li, "Adasyn: Adaptive synthetic sampling approach for imbalanced learning," in *2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence)*, Ieee, 2008, pp. 1322–1328.

[90] J. Beinecke and D. Heider, "Gaussian noise up-sampling is better suited than smote and adasyn for clinical decision making," *BioData Mining*, vol. 14, no. 1, p. 49, 2021.

[91] L. Breiman, "Random forests," *Machine learning*, vol. 45, pp. 5–32, 2001.

[92] S. Srinivasan, N. G. Harnett, L. Zhang, *et al.*, "Unravelling psychiatric heterogeneity and predicting suicide attempts in women with trauma-related dissociation using artificial intelligence," *European journal of psychotraumatology*, vol. 13, no. 2, p. 2 143 693, 2022.

[93] A. Vehtari, A. Gelman, and J. Gabry, "Practical bayesian model evaluation using leave-one-out cross-validation and waic," *Statistics and computing*, vol. 27, pp. 1413–1432, 2017.

[94] M. Hossin and M. N. Sulaiman, "A review on evaluation metrics for data classification evaluations," *International journal of data mining & knowledge management process*, vol. 5, no. 2, p. 1, 2015.

[95] Laboratory of Neuro Imaging, *LONI Atlases*, `https://loni.usc.edu/research/atlases`, Accessed: April 7, 2024.

[96] D. Thissen, L. Steinberg, and D. Kuang, "Quick and easy implementation of the benjamini-hochberg procedure for controlling the false positive rate in multiple comparisons," *Journal of educational and behavioral statistics*, vol. 27, no. 1, pp. 77–83, 2002.

[97] E. Langnes, M. H. Sneve, D. Sederevicius, I. K. Amlien, K. B. Walhovd, and A. M. Fjell, "Anterior and posterior hippocampus macro-and microstructure across the lifespan in relation to memory—a longitudinal study," *Hippocampus*, vol. 30, no. 7, pp. 678–692, 2020.

[98] S. H. Collin, B. Milivojevic, and C. F. Doeller, "Memory hierarchies map onto the hippocampal long axis in humans," *Nature neuroscience*, vol. 18, no. 11, pp. 1562–1564, 2015.

[99] Y. L. Rao, B. Ganaraja, B. Murlimanju, T. Joy, A. Krishnamurthy, and A. Agrawal, "Hippocampus and its involvement in alzheimer's disease: A review," *3 Biotech*, vol. 12, no. 2, p. 55, 2022.

[100] Q. Feng, J. Niu, L. Wang, *et al.*, "Comprehensive classification models based on amygdala radiomic features for alzheimer's disease and mild cognitive impairment," *Brain imaging and behavior*, vol. 15, pp. 2377–2386, 2021.

[101] S. Knafo, "Amygdala in alzheimer's disease," in *The Amygdala*, B. Ferry, Ed., Rijeka: IntechOpen, 2012, ch. 15. DOI: `10.5772/52804`. [Online]. Available: `https://doi.org/10.5772/52804`.

[102] M. S. Mega, J. L. Cummings, T. Fiorello, and J. Gornbein, "The spectrum of behavioral changes in alzheimer's disease," *Neurology*, vol. 46, no. 1, pp. 130–135, 1996.

[103] J. J. Knierim, "The hippocampus," *Current Biology*, vol. 25, no. 23, R1116–R1121, 2015.

[104] A. D. Korczyn and L. T. Grinberg, "Is alzheimer disease a disease?" *Nature Reviews Neurology*, pp. 1–7, 2024.

[105] H. Asai, N. Ohkawa, Y. Saitoh, *et al.*, "Pcdhβ deficiency affects hippocampal ca1 ensemble activity and contextual fear discrimination," *Molecular brain*, vol. 13, pp. 1–10, 2020.

[106] N. Insel and K. Takehara-Nishiuchi, "The cortical structure of consolidated memory: A hypothesis on the role of the cingulate–entorhinal cortical connection," *Neurobiology of Learning and Memory*, vol. 106, pp. 343–350, 2013.

[107] K. M. Igarashi, "Entorhinal cortex dysfunction in alzheimer's disease," *Trends in neurosciences*, vol. 46, no. 2, pp. 124–136, 2023.

[108] J. N. Giedd, A. Raznahan, K. L. Mills, and R. K. Lenroot, "Magnetic resonance imaging of male/female differences in human adolescent brain anatomy," *Biology of sex differences*, vol. 3, pp. 1–9, 2012.

[109] M. M. Mielke, P. Vemuri, and W. A. Rocca, "Clinical epidemiology of alzheimer's disease: Assessing sex and gender differences," *Clinical epidemiology*, pp. 37–48, 2014.

[110] A. Pfefferbaum, T. Rohlfing, M. J. Rosenbloom, W. Chu, I. M. Colrain, and E. V. Sullivan, "Variation in longitudinal trajectories of regional brain volumes of healthy men and women (ages 10 to 85 years) measured with atlas-based parcellation of mri," *Neuroimage*, vol. 65, pp. 176–193, 2013.

[111] L. L. Barnes, R. S. Wilson, J. L. Bienias, J. A. Schneider, D. A. Evans, and D. A. Bennett, "Sex differences in the clinical manifestations of alzheimer disease pathology," *Archives of general psychiatry*, vol. 62, no. 6, pp. 685–691, 2005.

[112] B. A. Ardekani, A. Convit, and A. H. Bachman, "Analysis of the miriad data shows sex differences in hippocampal atrophy progression," *Journal of Alzheimer's Disease*, vol. 50, no. 3, pp. 847–857, 2016.

[113] A. V. Witte, M. Savli, A. Holik, S. Kasper, and R. Lanzenberger, "Regional sex differences in grey matter volume are associated with sex hormones in the young adult human brain," *Neuroimage*, vol. 49, no. 2, pp. 1205–1212, 2010.

[114] E. Lentini, M. Kasahara, S. Arver, and I. Savic, "Sex differences in the human brain and the impact of sex chromosomes and sex hormones," *Cerebral cortex*, vol. 23, no. 10, pp. 2322–2336, 2013.

[115] A. E. Budson, *Why are women more likely to develop alzheimer's disease?* Harvard Health Publishing, URL: https://www.health.harvard.edu/blog/why-are-women-more-likely-to-develop-alzheimers-disease-202201202672, 2022.

[116] U.S. Census Bureau, *2023 national population projections tables: Main series*, Accessed: June 10, 2024, 2023. [Online]. Available: https://www.census.gov/data/tables/2023/demo/popproj/2023-summary-tables.html.

[117] N. I. on Aging (NIA), *Thinking about your risk for alzheimer's disease? five questions to consider*, https://www.nia.nih.gov/health/alzheimers-causes-and-risk-factors/thinking-about-your-risk-alzheimers-disease-five, Accessed: April 25, 2024, 2023.

[118] N. I. on Aging (NIA), *What causes alzheimer's disease?* https://www.nia.nih.gov/health/alzheimers-causes-and-risk-factors/what-causes-alzheimers-disease, Accessed: April 25, 2024, 2019.

[119] P. Bretsky, J. Buckwalter, T. Seeman, *et al.*, "Evidence for an interaction between apolipoprotein e genotype, gender, and alzheimer disease," *Alzheimer Disease & Associated Disorders*, vol. 13, no. 4, pp. 216–221, 1999.

[120] S. Subramaniapillai, A. Almey, M. N. Rajah, and G. Einstein, "Sex and gender differences in cognitive and brain reserve: Implications for alzheimer's disease in women," *Frontiers in Neuroendocrinology*, vol. 60, p. 100 879, 2021.

[121] L. A. Farrer, L. A. Cupples, J. L. Haines, *et al.*, "Effects of age, sex, and ethnicity on the association between apolipoprotein e genotype and alzheimer disease: A meta-analysis," *Jama*, vol. 278, no. 16, pp. 1349–1356, 1997.

[122] A. Association, *Medications for memory*, Webpage, Accessed: April 15, 2024, 2024. [Online]. Available: https://www.alz.org/alzheimers-dementia/treatments/medications-for-memory.

[123] M. Clinic, *Alzheimer's disease*, In-depth article, Accessed: April 15, 2024, 2024. [Online]. Available: https://www.mayoclinic.org/diseases-conditions/alzheimers-disease/in-depth/alzheimers/art-20048103.

[124] M. E. Hasselmo and J. M. Bower, "Acetylcholine and memory," *Trends in neurosciences*, vol. 16, no. 6, pp. 218–222, 1993.

[125] J. Grutzendler and J. C. Morris, "Cholinesterase inhibitors for alzheimer's disease," *Drugs*, vol. 61, pp. 41–52, 2001.

[126] J. T. Coyle, D. L. Price, and M. R. DeLong, "Alzheimer's disease: A disorder of cortical cholinergic innervation," *Science*, vol. 219, no. 4589, pp. 1184–1190, 1983.

[127] M. A. Rogawski and G. L. Wenk, "The neuropharmacological basis for the use of memantine in the treatment of alzheimer's disease," *CNS drug reviews*, vol. 9, no. 3, pp. 275–308, 2003.

[128] J. A. Harris, N. Devidze, L. Verret, *et al.*, "Transsynaptic progression of amyloid-β-induced neuronal dysfunction within the entorhinal-hippocampal network," *Neuron*, vol. 68, no. 3, pp. 428–441, 2010.

[129] S. Walsh, R. Merrick, E. Richard, S. Nurock, and C. Brayne, *Lecanemab for alzheimer's disease*, 2022.

[130] *Second death tied to experimental alzheimer's treatment*, Accessed: June 10, 2024, 2022. [Online]. Available: https://www.smithsonianmag.com/smart-news/second-death-tied-to-experimental-alzheimers-treatment-180981206/#:~:text=The%20second%20fatality%20was%20a,died%20a%20few%20days%20later..

[131] C. J. Swanson, Y. Zhang, S. Dhadda, *et al.*, "A randomized, double-blind, phase 2b proof-of-concept clinical trial in early alzheimer's disease with lecanemab, an anti-aβ protofibril antibody," *Alzheimer's research & therapy*, vol. 13, pp. 1–14, 2021.

[132] Y. Qiao, Y. Chi, Q. Zhang, and Y. Ma, "Safety and efficacy of lecanemab for alzheimer's disease: A systematic review and meta-analysis of randomized clinical trials," *Frontiers in aging neuroscience*, vol. 15, p. 1 169 499, 2023.

[133] S. H. Care, *Medications for alzheimer's disease*, Webpage, Accessed: April 15, 2024, 2024. [Online]. Available: https://stanfordhealthcare.org/medical-conditions/brain-and-nerves/alzheimers-disease/treatments/medications.html.

[134] U.S. Food and Drug Administration, *Fda approves treatment for adults with alzheimer's disease*, Accessed: 2024-07-08, 2024. [Online]. Available: https://www.fda.gov/drugs/news-events-human-drugs/fda-approves-treatment-adults-alzheimers-disease.

[135] D. Lowe, *Lecanemab and alzheimer's: More data*, Accessed: April 26, 2024, 2022. [Online]. Available: https://www.science.org/content/blog-post/lecanemab-and-alzheimer-s-more-data.

[136] U. P. Editors, "Alzheimer's disease and current treatment," *U.S. Pharmacist*, vol. 48, no. 6, pp. 22–24, Jun. 2023, Accessed: April 26, 2024. [Online]. Available: https://www.uspharmacist.com/article/alzheimers-disease-and-current-treatment.

[137] S. M. Fernandes, A. J. Mendes, P. F. Rodrigues, A. Conde, M. Rocha, and J. Leite, "Efficacy and safety of repetitive transcranial magnetic stimulation and transcranial direct current stimulation in memory deficits in patients with alzheimer's disease: Meta-analysis and systematic review," *International Journal of Clinical and Health Psychology*, vol. 24, no. 2, p. 100 452, 2024.

[138] S. B. Çelebi and B. G. Emiroğlu, "Leveraging deep learning for enhanced detection of alzheimer's disease through morphometric analysis of brain images.," *Traitement du Signal*, vol. 40, no. 4, 2023.

[139] J. Ashburner and K. J. Friston, "Voxel-based morphometry—the methods," *Neuroimage*, vol. 11, no. 6, pp. 805–821, 2000.

[140] S. M. Smith, M. Jenkinson, M. W. Woolrich, *et al.*, "Advances in functional and structural mr image analysis and implementation as fsl," *Neuroimage*, vol. 23, S208–S219, 2004.

[141] J. Lötjönen, R. Wolz, J. Koikkalainen, *et al.*, "Fast and robust extraction of hippocampus from mr images for diagnostics of alzheimer's disease," *Neuroimage*, vol. 56, no. 1, pp. 185–196, 2011.

[142] R. K. Olsen, D. J. Palombo, J. S. Rabin, B. Levine, J. D. Ryan, and R. S. Rosenbaum, "Volumetric analysis of medial temporal lobe subregions in developmental amnesia using high-resolution magnetic resonance imaging," *Hippocampus*, vol. 23, no. 10, pp. 855–860, 2013.

[143] D. Kennedy, C. Haselgrove, B. Fischl, *et al.*, "Harvard-oxford cortical and subcortical structural atlases," *Harvard Center for Morphometric Analysis*, 2016.