

Université du Québec
Institut national de la recherche scientifique
Centre Énergie Matériaux Télécommunications

**EXPLORING RESTING STATE EEG AND MRI FEATURES TO PREDICT
ALZHEIMER'S DISEASE SEVERITY LEVELS**

By
Belmir Jose de Jesus Junior

A thesis submitted in conformity with the requirements for the degree of
Master of Science, M.Sc.
in Telecommunications

Evaluation Committee

External evaluator	Prof. Sarah D. Power Memorial University
Internal evaluator	Prof. Marc Andre Gauthier INRS-EMT
Research supervisor	Prof. Tiago H. Falk

Acknowledgement

I would especially like to thank my supervisor, Dr. Tiago H. Falk, for sharing ideas, for offering availability to the clarification of doubts, for giving directions on how to proceed, and generally allowing me to do this work in such an enriching way. I want to express my sincere gratitude. I would like to extend my thanks to Prof. Sarah Power and Prof. Marc Andre Gauthier for accepting to serve as members of the evaluation committee, as well as the INRS-EMT staff and faculty members.

Next, I would like to thank all the MuSAE lab members for their incredible support, assistance, hard work, companionship and coffee time over the past few years. A huge thank you goes to Dr. Raymundo Cassani, who helped develop this work directly by consistently providing suggestions and comments.

Agradeço de forma grandiosa a minha familia, aos meus pais Belmir e Vera de Jesus juntamente as minhas irmãs Suzi e Jiovana Soares, os quais estão sempre a me motivar, inspirar e incentivar a nunca desistir dos meus sonhos. And I would like to include in this thanks those who became family by heart Austein McLoughlin, Diana and Fred Doern.

Agradeço de forma geral a todos os membros da minha família e aos amigos que se fizeram presente durante esse momento. Over the last few years, I have been helped and supported by so many. To those who have both, directly and indirectly, supported me during the master, thank you!

Abstract

There has been an increase in the number of people developing dementia, as global life expectancy also increases. Alzheimer's disease is still nowadays, an incurable and irreversible disorder. However, the development of therapies to decelerate the early stages of the disease has been a promising treatment option for people suffering from it. Therefore, current early diagnoses is still a challenge since the onset of symptoms is likely to happen twenty years after the physiological changes associated with the disease appears. Due to that, the diagnosis relies on invasive or expansive analyses, and so do the cognitive evaluations, mostly relying on "pencil and paper" tests, which require specialized healthcare professionals. The importance of measuring the patients' cognitive ability to evaluate possible treatments and even measure the disease's progression rate demand a more accessible and affordable technology. In this thesis, an attempt to develop an automated system that can predict the Mini-Mental State score of patients that have been diagnosed with Alzheimer's disease in its mild stage was investigated. The system was implemented from the exploration of features extracted from electroencephalography signals, and magnetic resonance imaging scans. Firstly, feature selection techniques were applied to selected the features that would best contribute to the task. Following, four different regression algorithms' performance utilizing features derived from each unique modality and then utilizing two modalities in combination. The preliminary experimental results demonstrated that the utilization of the complementary information carried by EEG measurements and MRI scans could be potentially useful for detecting the cognitive state of patients with Alzheimer's disease, and should be further explored.

Keywords Alzheimer's disease (AD), electroencephalography (EEG), magnetic resonance imaging (MRI), AD severity.

Résumé

Le nombre de personnes atteintes de démence a augmenté, car l'espérance de vie mondiale augmente également. La maladie d'Alzheimer est encore aujourd'hui une maladie incurable et irréversible. Cependant, le développement de thérapies visant à ralentir les premiers stades de la maladie a été une option de traitement prometteuse pour les personnes qui en souffrent. Par conséquent, les diagnostics précoces actuels constituent toujours un défi puisque l'apparition des symptômes est susceptible de se produire vingt ans après l'apparition des changements physiologiques associés à la maladie. De ce fait, le diagnostic repose sur des analyses invasives ou extensives, tout comme les évaluations cognitives, qui s'appuient principalement sur des tests "papier-crayon", nécessitant des professionnels de santé spécialisés. L'importance de mesurer la capacité cognitive des patients à évaluer les traitements possibles et même à mesurer le taux de progression de la maladie exige une technologie plus accessible et plus abordable. Dans cette thèse, une tentative de développement d'un système automatisé capable de prédire le score du mini-état mental des patients ayant reçu un diagnostic de maladie d'Alzheimer au stade léger a été étudiée. Le système a été mis en œuvre à partir de l'exploration des caractéristiques extraites des signaux d'électroencéphalographie et des scanners d'imagerie par résonance magnétique. Tout d'abord, des techniques de sélection des caractéristiques ont été appliquées pour choisir les caractéristiques qui contribueraient le mieux à la tâche. Ensuite, quatre algorithmes de régression différents ont été mis en œuvre en utilisant des caractéristiques dérivées de chaque modalité unique, puis en utilisant deux modalités en combinaison. Les résultats expérimentaux préliminaires ont démontré que l'utilisation des informations complémentaires véhiculées par les mesures EEG et les scanners IRM pourrait être potentiellement utile pour détecter l'état cognitif des patients atteints de la maladie d'Alzheimer, et qu'elle devrait être explorée plus avant.

Mots-clés la maladie d'Alzheimer, électroencéphalographie, imagerie par résonance magnétique, la gravité de la maladie d'Alzheimer.

Contents

Acknowledgement	iii
Abstract	v
Résumé	vii
Contents	ix
List of figures	xiii
List of tables	xv
List of abbreviations	xvii
Synthèse	1
0.1 Introduction	1
0.1.1 Les phases de la maladie d'Alzheimer	2
0.1.2 Diagnostic et traitement de la maladie d'Alzheimer	3
0.1.3 Gravité de la maladie d'Alzheimer	4
0.1.4 Objectifs de la thèse	5
0.2 Biomarqueurs et système de monitorage de la gravité de la MA	5
0.2.1 IRM et maladie d'Alzheimer	5
0.2.2 EEG et maladie d'Alzheimer	6
0.2.3 Système multimodal	8
0.2.4 Aperçu du système de prédition automatisé	8
0.3 Matériaux et méthodes	9
0.3.1 Participants	9
0.3.2 Collecte de données	10
0.3.3 Prétraitement	10
0.3.4 Extraction des caractéristiques	11
0.3.5 Sélection des caractéristiques	12
0.3.6 Regroupement des principales caractéristiques	13
0.3.7 Partitionnement des données	13
0.3.8 Scénarios de régression	14
0.3.9 Mesure d'évaluation	14
0.4 Résultats et discussions	15
0.4.1 Analyse statistiques des caractéristiques	15
0.4.2 Prévision de la gravité de l'AD	16

0.4.3	Caractéristiques pertinentes	17
0.5	Conclusions	19
0.5.1	Limites de l'étude	20
0.5.2	Travaux futurs	21
1	Introduction	23
1.1	Phases of Alzheimer's disease	25
1.2	Diagnosis and treatment of Alzheimer's disease	26
1.3	Biomarkers for Alzheimer's disease	27
1.3.1	Cerebrospinal fluid (CSF)	27
1.3.2	Positron emissions tomography (PET) scans	27
1.3.3	Magnetic resonance imaging (MRI)	28
1.3.4	Electroencephalography (EEG)	28
1.3.5	Magnetoencephalography (MEG)	28
1.4	Alzheimer's disease severity	28
1.5	Thesis objective	29
1.6	Publications	30
1.7	Thesis organization	30
2	Biomarkers and AD severity monitoring system	31
2.1	Magnetic resonance imaging	31
2.2	MRI and Alzheimer's disease	32
2.3	Electroencephalography	33
2.4	EEG and Alzheimer's disease	33
2.4.1	Slowing of EEG	34
2.4.2	Reduced complexity	34
2.4.3	Perturbation in EEG synchrony	34
2.4.4	Neuromodulatory deficit	34
2.5	Multimodal systems	35
2.5.1	Data-level fusion	35
2.5.2	Feature-level fusion	35
2.5.3	Decision-level fusion	36
2.5.4	Hybrid fusion	36
2.6	Automated prediction system overview	36
3	Materials and methods	39
3.1	Participants	39
3.2	Data collection	40
3.3	Pre-processing	40
3.4	Feature extraction	41
3.4.1	EEG features	41
3.4.2	MRI feature extraction	46
3.5	Feature selection	49
3.5.1	Correlation-based ranking	49
3.5.2	Minimum redundancy maximum relevance, MRMR	50
3.5.3	Top-features grouping	50
3.6	Data partitioning	51
3.7	Classification/Regression	52

3.7.1	Support vector machine	52
3.7.2	Random forest	52
3.7.3	K-nearest neighbour	52
3.8	Evaluation metric	53
4	Results and discussion	55
4.1	Statistical analyses	55
4.2	AD severity prediction	58
4.3	Consistently selected features	61
5	Conclusion	69
5.1	Study limitations	70
5.2	Future work	70
	Bibliography	71

List of figures

2.1	Workflow of the prediction system.	37
3.1	(a) Cognision proprietary COGNISION TM system. (b) EEG electrode placement for the AD dataset. Image taken from [21].	40
3.2	(a) Signal processing steps involved in computation of amplitude modulation rate-of-change features. (b) Bandpass filtered EEG signals and their time envelopes. Adapted from [38]	43
3.3	Regions identified in the modulation power domain. Image taken from [20].	44
3.4	Distributions of the feature selection-train-test data partitions used.	51
4.1	Relationship among the EEG and MRI extracted features.	58
4.2	Distribution of the RMSE scores according to regressors used.	60
4.3	Distribution of the RMSE scores according to the features selection techniques.	60
4.4	Distribution of the RMSE scores according to the groups of features.	61

List of tables

3.1	Demographics for the patients used in this study	39
3.2	Summary of the features computed from EEG signals.	45
3.3	Summary of the MRI computed features.	48
3.4	Top-feature grouping classes.	51
4.1	Ranking of the EEG features correlated with MMSE scores. All the values presented a p-value smaller than 0.05.	56
4.2	Ranking of the MRI features correlated with MMSE scores. All the values presented a p-value smaller than 0.05.	57
4.3	Performance comparison across regressors, feature selection algorithms, and feature groups.	59
4.4	Consistently Selected Features from Group 1 - EEG Features.	62
4.5	Consistently Selected Features from Group 2 - MRI Features.	64
4.6	Consistently Selected Features from Group 3 - EEG and MRI Features.	65
4.7	Distribution of the selected features that were selected in multiple runs for EEG groups.	66
4.8	Distribution of the selected features that were selected in multiple runs for MRI groups.	67
4.9	Comparative regression performance from models formed by the top features from Tables 4.4, 4.5 and 4.6	67

List of abbreviations

AD Alzheimer's disease

CSF Cerebrospinal fluid

MCI Mild cognitive impairment

PET Positron emissions tomography

EEG Electroencephalography

MRI Magnetic resonance imaging

fMRI Functional magnetic resonance imaging

MEG Magnetoencephalography

MMSE Mini-mental state examination

NMR Nuclear magnetic resonance

PCC Posterior cingulate cortex

FIR Finite impulse response filter

wICA Wavelet-enhanced independent component analysis

IC Independents components

MSC Magnitude squared coherence

STD Standard deviation

CV Coefficient of variation

CT Cortical thickness

ROI Region-of-interest

ASEG Automatic subcortical segmentation

WM White matter

eTIV Estimated total intracranial volume

MRMR Minimum redundancy maximum relevance

SVM Support vector machines

RF Random forest

KNN K-nearest neighbours

RMSE Root mean squared error

Synthèse

0.1 Introduction

Les chercheurs définissent la maladie d'Alzheimer (MA) comme un trouble dégénératif progressif du cerveau, qui est considéré comme la cause la plus fréquente de la démence. La caractérisation de la démence est une conjonction de symptômes tels que le déclin fonctionnel de la mémoire, du langage, de la résolution de problèmes et d'autres aptitudes cognitives qui peuvent avoir un impact direct sur la qualité de vie de la personne. Les symptômes apparaissent en raison de la mort ou de l'endommagement des cellules nerveuses du cerveau (neurones), qui pourrait commencer vingt ans ou plus avant la manifestation des symptômes [3]. La MA est une maladie neurodégénérative qui touche majoritairement les personnes de plus de 65 ans et qui est malheureusement encore aujourd'hui incurable [2].

Des études ont montré qu'en 2016, une population de 26,6 millions de sujets vivaient avec la MA dans le monde entier. En outre, les modèles prévoient que d'ici 2050, ce nombre sera multiplié par quatre, ce qui signifie qu'une personne sur 85 aura développé la MA. De manière substantielle, à mesure que l'espérance de vie de la population augmente, le nombre de personnes âgées augmente également. Un moyen d'éviter que le monde ne soit confronté à l'expansion mondiale actuelle de MA pourrait être les progrès encourus dans la planification préventive et des thérapies visant à ralentir la progression de la MA, puisque l'on estime que 43% des cas en 2050 nécessiteront des niveaux de soins élevés [14].

Les personnes souffrant de la MA peuvent développer divers symptômes, dont l'intensité peut varier selon le stade de la maladie. Certains symptômes de la maladie comprennent généralement un déclin intellectuel et comportemental rapide, qui est généralement associé à des difficultés de

mémoire à court terme. Comme indiqué dans [2], les symptômes se manifestent généralement selon quatre modalités principales. Les symptômes physiologiques comprennent des modifications du tissu cérébral dues à la présence de plaques et d'enchevêtrements, une hypoperfusion cérébrale et une atrophie dans les zones corticales et hippocampiques du cerveau. La dépression, les changements d'humeur et l'apathie, à leur tour, sont quelques exemples des symptômes psychologiques auxquels sont confrontés les patients atteints de la MA. Les symptômes cognitifs consistent en des troubles de la mémoire, des obstacles à la résolution des problèmes quotidiens et des problèmes de communication. Enfin, en conséquence des symptômes mentionnés précédemment, la personne atteinte peut présenter un comportement incohérent.

0.1.1 Les phases de la maladie d'Alzheimer

La littérature rapporte que les personnes atteintes de la MA passent par trois phases : pré-clinique, déficience cognitive légère et démence due à la MA [18, 2]. La maladie commence bien avant que l'un des symptômes n'apparaisse dans la première phase appelée MA préclinique. En raison de l'absence de symptômes, il est difficile de détecter la maladie à ce stade. Cependant, des modifications du cerveau, du sang et du liquide céphalo-rachidien (LCR) peuvent déjà apparaître.

Le deuxième stade de la maladie est appelé déficience cognitive légère (DCL). Au cours de cette phase, les problèmes liés à la capacité de réflexion, comme les pertes de mémoire, la difficulté à choisir le bon vocabulaire et l'emplacement géographique des lieux familiers deviennent apparents. Même si les symptômes sont perceptibles pour les patients eux-mêmes et/ou leur entourage, l'indépendance des individus dans les activités de la vie quotidienne n'est pas affectée à ce stade. Cependant, toutes les personnes qui reçoivent un diagnostic de DCL ne développeront pas systématiquement une démence due à la MA. On ne sait toujours pas pourquoi certaines personnes développent une démence alors que d'autres n'en développent pas.

Au dernier stade de la maladie, les patients développent une démence due à la MA. À ce stade, les troubles de la mémoire, de la pensée et du comportement sont très visibles, ce qui affecte considérablement les activités quotidiennes du patient. Deux sous-étapes différentes peuvent découler de la démence due à la MA, à savoir la MA modérée ou moyenne et la MA grave ou tardive. Le premier fait référence à la phase la plus prolongée de la maladie, où l'exécution des tâches de la vie quotidienne devient un problème pour les individus, ce qui entraîne de la frustration et de la

colère. La seconde est le stade final où les capacités cognitives et la mémoire des patients sont endommagées, ce qui nécessite une assistance personnelle 24 heures, puisque les mouvements et les capacités physiques des patients sont également affectés.

0.1.2 Diagnostic et traitement de la maladie d'Alzheimer

Actuellement, un diagnostic précis de la MA n'est possible qu'après la réalisation d'un examen structurel du cerveau, après le décès des patients. Actuellement, un diagnostic précis de la MA est possible grâce à un examen structurel post-mortem du cerveau. Traditionnellement, le diagnostic repose principalement sur une évaluation du déclin mental basée sur des tests papier-crayon, tels que l'évaluation cognitive de Montréal (MoCA) [76] et/ou le mini-examen de l'état mental (MMSE) [11]. Avec le MMSE, par exemple, un clinicien professionnel pose au patient plusieurs questions sur ses compétences mentales quotidiennes pour évaluer son état cognitif. Le participant reçoit ensuite un score basé sur ses réponses, avec un maximum de 30 points [45]. Les scores de 20 à 24 indiquent une démence légère, tandis que les scores de 13 à 20 suggèrent une démence modérée, et ceux de moins de 12 indiquent une démence grave.

Ces tests visent à mesurer les capacités cognitives et fonctionnelles des patients, qui surviennent généralement lorsque la maladie a déjà causé des dommages cérébraux notables. C'est pourquoi la recherche se développe pour trouver des biomarqueurs de la maladie avant l'apparition des symptômes. On sait que les interventions thérapeutiques sont plus efficaces lorsqu'elles sont administrées à des stades précoce de la maladie, avant que la perte neuronale ne se produise [55], ce qui rend la détection précoce cruciale. Au cours des deux dernières décennies, un certain nombre de biomarqueurs ont été explorés, notamment ceux basés sur la neuroimagerie structurelle (par exemple, le volume de l'hippocampe à partir de l'imagerie par résonance magnétique) [1], les fluides cérébraux de la colonne vertébrale (par exemple, niveaux tau et bêta-amyloïde) [16], les échantillons de sang et d'urine [74], l'électroencéphalographie [18], et plus récemment, le profilage du risque génétique [99].

La maladie d'Alzheimer étant encore une maladie incurable, le traitement est généralement multiforme et repose sur la gestion des symptômes et le ralentissement du déclin clinique. Pour un traitement favorable, la première étape est la prise de conscience des symptômes, suivie d'un diagnostic correct. À ce stade, le patient et les soignants doivent être informés sur la maladie,

l'évolution attendue, les options de traitement et les attentes concernant le développement de la maladie. Des approches non pharmacologiques peuvent être suggérées en fonction des besoins, des priorités et des ressources du patient et des soignants, afin d'atténuer les symptômes tels que l'agitation, l'apathie, les délires, la désorientation, la résistance aux soins, entre autres. Par la suite, la mise en place d'un traitement pharmacologique adapté peut être nécessaire. À long terme, les plans de soins doivent être ajustés et modifiés en fonction des changements et des conditions du patient ainsi que des soignants [5].

0.1.3 Gravité de la maladie d'Alzheimer

Lorsqu'ils reçoivent un diagnostic de MA, les patients et leurs familles commencent à se poser des questions sur la progression du déclin cognitif. Malheureusement, le rythme de progression de la maladie d'Alzheimer varie considérablement et, en moyenne, les personnes atteintes peuvent vivre entre trois et vingt ans ou plus après le diagnostic, ce qui laisse aux cliniciens la tâche difficile de surveiller la progression de la maladie [33]. En outre, à mesure que de nouveaux traitements, médicaments et interventions apparaissent, le suivi de la progression et/ou du ralentissement de la maladie deviendra crucial pour évaluer les avantages des différentes interventions [101]. Cela est particulièrement difficile dans le cas d'une détection précoce, car les symptômes peuvent ne pas être aussi prononcés.

Si plusieurs études ont montré les avantages des biomarqueurs pour la détection précoce de la MA, peu d'entre elles ont exploré leur utilité pour le suivi de la progression de la maladie. Dans [55], par exemple, plusieurs biomarqueurs de la base de données de l'initiative de neuroimagerie de la maladie d'Alzheimer (ADNI) ont été explorés et les niveaux de volume, de tau et de bêta-amyloïde de l'hippocampe se sont révélés prometteurs. Dans [36], à son tour, l'atrophie de l'hippocampe et l'épaisseur du cortex se sont révélées être des prédicteurs utiles de la progression de la déficience cognitive légère vers la MA. Plus récemment, dans [18], il a été montré que seules quelques études ont exploré les effets de la progression de la maladie sur l'EEG. En particulier, des travaux antérieurs ont montré que des changements dans des sous-bandes de fréquence spécifiques pouvaient être observés avec la progression de la maladie (par exemple, [64, 65]). Plus récemment, des patchs d'énergie spectro-temporelle ont montré qu'ils offraient une meilleure précision (par rapport aux puissances de sous-bandes classiques) lors de la prévision des niveaux de gravité de la maladie [19].

0.1.4 Objectifs de la thèse

Comme la maladie d’Alzheimer se manifeste sous différentes facettes, on s’attend à ce qu’un système de biomarqueurs multimodaux soit nécessaire pour surveiller avec précision l’évolution de la maladie d’Alzheimer. En fait, certaines études suggèrent qu’un seul biomarqueur pourrait ne pas fournir suffisamment d’informations pour déterminer avec précision la gravité de la maladie [50]. Alors que des biomarqueurs multimodaux ont été étudiés pour la détection de la MA et de la DCL [102], la plupart ont fait appel à des outils de neuroimagerie structurelle [37] et seule une poignée a exploré la fusion de la neuroimagerie structurelle avec l’électro-neurophysiologie [18]. En outre, des travaux limités ont été explorés pour les systèmes multimodaux de détection de la gravité de la MA [73]. Le présent document vise à combler cette lacune.

Plus précisément, cette thèse explore le développement d’un corrélat multimodal du niveau de gravité de la MA pour les patients aux premiers stades de la maladie (scores MMSE entre 21 et 26). Nous explorons la fusion de plusieurs caractéristiques structurelles, de l’épaisseur corticale et de la substance blanche extraites de l’IRM, ainsi que des caractéristiques EEG classiques avec celles récemment développées, telles que les caractéristiques du patch d’énergie spectro-temporelle [19]. Les expériences menées sur 89 patients montrent l’utilité des modèles multimodaux de niveau de gravité de la MA pour les patients aux premiers stades de la maladie.

0.2 Biomarqueurs et système de monitorage de la gravité de la MA

0.2.1 IRM et maladie d’Alzheimer

L’IRM structurelle du cerveau a été largement étudiée en tant qu’outil de neuroimagerie pour aider à diagnostiquer la MA [2]. Les progrès de la recherche ont permis de développer des méthodes automatisées qui peuvent effectuer une segmentation sous-corticale des régions du cerveau et déterminer de mesurer l’épaisseur et du volume cortical, mesures essentielles qui aident à détecter les biomarqueurs de la MA [61]. L’IRM s’est également révélée être un outil potentiel pour le monitorage de la progression de la maladie par des mesures volumétriques [52].

Plusieurs études ont montré la capacité de l'IRM structurelle à évaluer l'atrophie de certaines zones du cerveau. La réduction du volume de l'hippocampe et du cortex entorhinal sont les marqueurs les plus persistants rencontrés dans les examens IRM des patients atteints de la MA. L'atrophie de l'hippocampe, cependant, n'est pas spécifique de la MA et a également été associée à d'autres pathologies, telles que la schizophrénie et la dépression [52], d'où la nécessité de systèmes multimodaux.

La littérature rapporte également une diminution du volume de matière grise dans le gyrus parahippocampique, dans l'amygdale, dans le cortex d'association postérieur et dans les noyaux sous-corticaux comme marqueurs de la MA [52]. Les mesures du volume et de l'épaisseur corticale du précuneus et du cortex cingulaire postérieur (PCC), du lobe pariétal, du lobe temporal et frontal ont également été associées à la neuropathologie de la MA [61].

De nombreux chercheurs ont engagé leurs efforts dans le développement d'un système de diagnostic automatique des maladies d'Alzheimer, car il serait un assistant important pour les cliniciens et les patients. Certains des travaux qui ont utilisé l'IRM pour poursuivre la production d'un tel système sont [81, 7]

0.2.2 EEG et maladie d'Alzheimer

L'EEG s'est révélé être une ressource prometteuse pour le diagnostic et l'investigation de la MA en raison de ses avantages en termes de portabilité, de prix et de non-invasivité par rapport aux autres techniques. Encouragés par ses avantages, les chercheurs ont étudié de quelle manière la MA et sa progression pourraient affecter les signaux EEG. Une récente revue systématique de la littérature a identifié quatre effets récurrents de la MA sur les enregistrements EEG au repos [18].

Comme les approches basées sur l'IRM pour le diagnostic clinique de la MA, l'EEG pourrait offrir des avantages similaires, tout en fournissant une application de diagnostic plus automatique. Des exemples de travaux qui ont utilisé l'EEG dans le développement d'un système de diagnostic automatique sont [21, 38].

Ralentissement de l'EEG

Un effet significatif observable dans les signaux EEG est associé à une puissance plus élevée dans les bandes de basse fréquence, telles que delta et thêta, et à une puissance plus faible dans les bandes de haute fréquence, telles que alpha et bêta, ce qui suggère un ralentissement de l'EEG par la MA [9]. De plus, ce décalage fréquentiel semble être proportionnel à la progression de la MA [18].

Complexité réduite

Les signaux EEG recueillis auprès des patients atteints de la MA semblent moins complexes que ceux obtenus auprès des mêmes sujets du groupe d'âge, ce qui indique que les signaux EEG des patients atteints de la MA sont composés de formes d'ondes plus similaires, tandis que les signaux EEG des sujets en bonne santé sont plus complexes ou "caotique". On pense que ce changement dans la dynamique neuronale avec la MA résulte de la perte accrue de neurones, entraînant ainsi une diminution de l'interaction neuronale et, par conséquent, de la complexité de l'EEG [57].

Perturbation de la synchronisation EEG

De nombreuses études ont montré une diminution de la synchronisation de l'EEG chez les patients atteints de la MA, ce qui suggère une diminution des schémas attendus dans différentes régions du cerveau et une perte de connectivité entre certaines régions corticales. La cause de cette perturbation dans la synchronisation de l'EEG n'est pas encore entièrement comprise [31].

Déficit neuromodulatoire

L'analyse de la modulation d'amplitude a été appliquée pour mesurer quantitativement l'activité neuromodulatoire des signaux EEG au repos. Des méthodes récentes ont montré des déficits neuromodulateurs et une dynamique de fréquence croisée réduite chez les patients atteints de la MA par rapport aux groupes de contrôle appariés par âge [38, 46].

0.2.3 Système multimodal

Un système multimodal correspond à un système de diagnostic/suivi de niveau de gravité qui fonctionne sur la base de la combinaison de deux ou plusieurs modalités d'entrée de l'utilisateur pour produire le résultat souhaité. Dans [102], il est souligné que la majorité des systèmes de diagnostic de la MA signalés reposent encore sur une modalité individuelle. Il est démontré que les biomarqueurs dérivés de différentes modalités peuvent fournir des informations complémentaires qui peuvent améliorer la robustesse et la précision du système.

Pour fonctionner comme un système multimodal, la fusion de différentes modalités d'entrée doit être effectuée. Ceci peut être réalisé en utilisant trois méthodes différentes: la fusion au niveau des données, des caractéristiques et des décisions. La fusion n'est pas triviale, car différentes modalités peuvent avoir été collectées à des taux d'échantillonnage différents, avec des capteurs et des résolutions variables (par exemple, séries temporelles 1-d, images 2-d ou images/tensors 3-d), et/ou fournissant des informations redondantes [34].

0.2.4 Aperçu du système de prédiction automatisé

Comme mentionné précédemment, les objectifs de cette thèse sont doubles: (1) construire un système multimodal qui combine les fonctions EEG et IRM pour une meilleure performance, et (2) explorer l'utilisation d'une nouvelle fonction de patch de spectre de modulation EEG pour la mesure du niveau de gravité de la MA. La figure 2.1 illustre un schéma fonctionnel de haut niveau d'une solution de prédiction automatisée. L'image montre un système unimodal, mais qui peut facilement être étendu à plusieurs modalités.

Le premier module consiste en l'acquisition de signaux. C'est l'étape au cours de laquelle les données brutes sont collectées à partir de différents types de capteurs, stockées et mises à disposition pour visualisation et analyse. Pour cette thèse, l'acquisition de signaux comprend la mesure des signaux EEG pendant un protocole au repos et l'acquisition de scanners IRM du cerveau. Ensuite, comme la qualité des données acquises n'est pas toujours élevée et qu'elles peuvent être contaminées par des artefacts, une étape de prétraitement est appliquée. Ici, l'objectif principal est d'améliorer la qualité des données brutes acquises via des algorithmes d'élimination des artefacts et/ou de suppression du bruit.

Une fois que les données ont été améliorées et les artefacts indésirables éliminés, l'étape suivante correspond à l'extraction des caractéristiques, où les propriétés ou informations importantes des signaux mesurés peuvent être extraites. Généralement, les caractéristiques sont d'une résolution beaucoup plus faible que les modalités brutes originales (par exemple, séries temporelles ou image 2-d) et transmettent des informations importantes pour la tâche de classification/régression en aval. Généralement, on ne sait pas à l'avance quelles caractéristiques fourniront la meilleure discrimination entre les classes, c'est pourquoi de nombreuses caractéristiques "candidates" sont extraites. Avec cette approche, cependant, de nombreuses caractéristiques peuvent être corrélées les unes avec les autres, certaines peuvent être plus sensibles au bruit que d'autres, et certaines peuvent ne pas fournir de pouvoir de discrimination, donc les inclure dans le système de prédiction peut être désavantageux. À cette fin, la sélection des caractéristiques est utilisée de manière à ce que les caractéristiques les plus pertinentes pour la tâche à accomplir soient trouvées et conservées. En réduisant le nombre de caractéristiques, les systèmes deviennent généralement moins sensibles aux biais des bases de données et sont moins sujets à la "malédiction de la dimensionnalité" et surapprentissage.

Une fois que les caractéristiques les plus pertinentes ont été trouvées, la dernière étape du système de prédiction comprend l'étape de classification ou de régression. C'est à ce stade que les modèles sont formés pour apprendre à établir des correspondances entre les caractéristiques sélectionnées en premier lieu et les étiquettes fournies (en supposant qu'un apprentissage supervisé est utilisé). La classification est généralement utilisée dans les systèmes de diagnostic, où des étiquettes discrètes telles que "contrôle" ou "MA" sont utilisées. La régression, quant à elle, vise à prédire une variable de résultat continue, telle que le score MMSE ; celles-ci sont applicables aux systèmes de monitorage de l'évolution de la maladie.

0.3 Matériaux et méthodes

0.3.1 Participants

Les données utilisées ici ont été collectées à partir d'un essai clinique multi-site explorant l'utilisation de biomarqueurs du potentiel événementiel (PRE) de la maladie d'Alzheimer [26]. Les données ont été initialement recueillies auprès de 103 sujets atteints d'une MA légère probable et de 101 témoins sains. L'étude a été approuvée par les comités d'examen institutionnels de chaque

site d'essai et le consentement de chaque participant a été obtenu. Le lecteur intéressé est invité à consulter [26] pour plus de détails sur les critères d'inclusion et d'exclusion, ainsi que sur tous les tests et critères effectués pour les diagnostics.

Ici, nous sommes particulièrement intéressés par les données de la population cible où les EPR, les EEG au repos et les scanners IRM structurels ont été recueillis. Après avoir soigneusement vérifié que la qualité des scans était suffisante pour la segmentation/parcellation de l'IRM, 89 participants au total étaient disponibles pour une analyse plus approfondie. Le tableau 3.1 résume les données démographiques de ces 89 participants. Tous les patients ont présenté un score MMSE entre 21 et 26.

0.3.2 Collecte de données

Les signaux EEG ont été acquis à l'aide du dispositif COGNITIONTM [17], un dispositif à sept canaux fonctionnant à une fréquence d'échantillonnage de 125 Hz. Des électrodes référentielles biauriculaires ont également été fixées pendant la procédure de collecte. Pour chacun des participants, nous nous intéressons à l'EEG collecté pendant une période de repos de 3 minutes en état d'éveil, les yeux ouverts. La figure ?? montre le placement des électrodes EEG situées aux emplacements F3, Fz, F4, Cz, P3, Pz et P4. Motivé par [38], deux signaux bipolaires inter-hémisphériques virtuels sont également utilisés, à savoir F3-F4 et P3-P4. Les données d'IRM avec volumétrie, à leur tour, ont été obtenues à partir de scans structurels pondérés T1. Les scans n'ont présenté aucun mouvement ou artefacts de segmentation ultérieurs.

0.3.3 Prétraitement

Les données EEG ont d'abord été filtrées par un filtre passe-bande à réponse impulsionale finie (FIR) à phase nulle avec une largeur de bande de 0,5 à 45 Hz ; c'est ce qu'on appelle "l'EEG brut". Comme les signaux EEG sont souvent contaminés par des artefacts oculaires, musculaires, de mouvement et/ou de lignes électriques [30], des algorithmes d'élimination des artefacts sont également employés. Ici, sur la base des informations fournies par [22], la méthode d'analyse en analyse d'ondelette en composante indépendante (wICA) a été utilisée. L'algorithme wICA décompose les signaux EEG en composantes indépendantes (CI), qui sont ensuite transformées à

l'aide d'une transformée en ondelette. Les coefficients d'ondelette obtenus sont ensuite soumis à un seuil pour séparer les composantes liées à l'activité neuronale de celles liées aux artefacts. Une fois les composantes de l'activité neuronale sélectionnées, une transformation inverse en ondelette est appliquée et les CI corrigés sont ensuite projetés dans le temps pour obtenir un signal amélioré [25] ; celui-ci est appelé "wICA EEG".

Les scans IRM pondérés T1, à leur tour, ont été traités par Freesurfer (<https://surfer.nmr.mgh.harvard.edu>), une plateforme ouverte universellement utilisée pour traiter et analyser les scans IRM structurels du cerveau qui a été validée par les chercheurs en IRM [42]. Les images capturées par le scanner IRM ont été prétraitées par la commande par défaut "reconall" fournie dans le pipeline de freesurfer afin de mettre en œuvre la correction et la normalisation des mouvements.

0.3.4 Extraction des caractéristiques

Extraction des caractéristiques de l'EEG

Avant l'extraction des caractéristiques, les signaux EEG d'une durée de 3 minutes étaient divisés en fenêtres (appelées epoch) d'une durée de 8 secondes avec des décalages d'une seconde entre les époques consécutives. Ensuite, les caractéristiques ont été calculées pour chaque epoch. Un certain nombre de caractéristiques ont été extraites, comme la puissance, la cohérence spectrale, taux de modulation de l'amplitude et caractéristiques des "patchs" de fréquence de modulation. Au total, 286 caractéristiques ont été extraites de chacune des époques. Le tableau 3.2 résume les caractéristiques calculées pour chaque epoch EEG. Comme mentionné ci-dessus, les caractéristiques EEG sont calculées par epoch. Les caractéristiques IRM, quant à elles, sont disponibles par sujet. Ainsi, afin de fusionner les caractéristiques EEG et IRM, nous devons regrouper toutes les caractéristiques EEG en une caractéristique finale par sujet. De plus des propriétés statistiques de chaque type de caractéristiques ont aussi été extraites, permettant d'évaluer leur évolution dans le temps. Plus précisément, la moyenne, l'écart-type, le coefficient de variation, la médiane, l'asymétrie et l'aplatissement ont été obtenus pour chacun des quatre types de caractéristiques EEG (puissance spectrale, cohérence, modulation d'amplitude et caractéristiques des "patchs" de fréquence de modulation) [32]. Au final, un total de 1716 caractéristiques EEG correspondant aux 286 caractéristiques multipliées par les 6 fonctions statistiques.

Extraction des caractéristiques de l'IRM

Les caractéristiques de l'IRM ont été extraites à l'aide du logiciel Freesurfer. Les volumes sous-corticaux, l'épaisseur corticale (CT) et la surface ont été extraits des régions d'intérêt (ROI) respectives. Les caractéristiques ont été extraites par le groupe du Dr William J. McGeown de l'Université de Strathclyde dans le cadre d'une collaboration avec Cognision [75].

Le tableau 3.3 résume les 285 caractéristiques calculées pour chacun des scans IRM. Par la suite, toutes les caractéristiques calculées ont été soumises à un processus de normalisation. La normalisation a été utilisée pour mettre à l'échelle toutes les caractéristiques dans une plage comparable de la valeur de -1 à 1 et éviter les problèmes futurs avec les modèles de régression qui pourraient dépendre des calculs de distance.

0.3.5 Sélection des caractéristiques

À la fin de la procédure d'extraction des caractéristiques, un total de 1716 caractéristiques ont été calculées à partir des signaux EEG et un total de 285 caractéristiques ont été calculées à partir des scans IRM. Bien que notre objectif principal soit d'explorer les avantages d'un système multimodal, l'un des buts exploratoires que nous poursuivons est également d'évaluer les avantages d'une modalité unique par rapport à l'autre, en particulier les avantages des nouvelles caractéristiques de patchs de spectre de modulation proposées [20]. Dans le système multimodal, nous utilisons la technique de fusion de caractéristiques pour combiner les modalités. Dans nos expériences, la fusion de caractéristiques est réalisée en concaténant les matrices de caractéristiques en une matrice de caractéristiques finale plus grande.

Compte tenu du grand nombre de caractéristiques et de la taille limitée de l'ensemble de données disponibles, il est nécessaire de sélectionner les caractéristiques les plus pertinentes pour éviter le surapprentissage [59]. La sélection ou le classement des caractéristiques vise à classer les caractéristiques disponibles en fonction de leur impact potentiel sur la tâche de classification en aval [15]. De cette façon, les caractéristiques non pertinentes et/ou redondantes sont supprimées.

De nombreuses techniques de sélection des caractéristiques ont été rapportées dans la littérature et il n'existe pas de consensus sur la technique idéale pour une application spécifique, c'est pourquoi différentes méthodes sont couramment explorées. Différentes techniques de filtrage univariées pour

la sélection des caractéristiques sont donc étudiées. Les techniques de filtrage n'utilisent aucun modèle d'apprentissage pendant le processus de sélection des caractéristiques, le classement est donc basé uniquement sur les caractéristiques des données [6]. Combinée à une approche univariée, l'étape de sélection des caractéristiques évalue une par une chaque caractéristique de l'ensemble de données en déterminant leur relation avec la variable dépendante, étiquette à prédire ; dans notre cas, les scores du MMSE. Au final, les caractéristiques sont classées en fonction de cette relation et les caractéristiques ayant obtenu les meilleurs scores sont conservées tandis que les autres sont supprimées. Dans cette étude, trois techniques différentes de filtrage des caractéristiques ont été explorées, le classement basé sur la corrélation (Pearson and Spearman) et revelance minimun et la redondance maximun, MRM.

0.3.6 Regroupement des principales caractéristiques

Dans l'optique de maintenir une dimensionnalité similaire avec celle de la literature (par exemple, [38, 47, 23]), uniquement 24 caractéristiques les plus pertinentes ont été selectionnées pour chaque modalités. Ici, nous appelons "groupe 1" les 24 principales caractéristiques sélectionnées pour la modalité EEG. Le groupe 2, quant à lui, correspond aux 24 principales caractéristiques de l'IRM. Comme nous sommes particulièrement intéressés par les systèmes multimodaux, deux autres stratégies de fusion de caractéristiques sont explorées. Premièrement, nous appliquons la sélection des caractéristiques aux ensembles de caractéristiques EEG-IRM fusionnés et nous conservons les 24 principales caractéristiques de cet ensemble combiné. Des travaux antérieurs ont suggéré une certaine corrélation entre les caractéristiques d'EEG et d'IRM, ce qui fait que la sélection des caractéristiques de l'ensemble fusionné peut aider à éliminer les caractéristiques redondantes de différentes modalités. Enfin, nous combinons les 12 premières caractéristiques EEG et les 12 premières caractéristiques IRM et appelons cette caractéristique le groupe 4. Le tableau 3.4 résume les quatre principaux groupes de caractéristiques examinés ici.

0.3.7 Partitionnement des données

Sur les données disponibles des 89 patients, 25 % ont été exclusivement consacrés à la procédure de sélection des caractéristiques, tandis que les 75 % restants ont été utilisés pour la formation et le test des modèles de régression. L'utilisation d'ensembles disjoints pour la sélection des car-

actéristiques et l’entraînement des classificateurs réduit tout biais indésirable dans les chiffres de performance rapportés. Le partitionnement des données a été effectué de manière aléatoire tout en gardant la distribution des scores du MMSE proportionnelle à l’ensemble des données. Le partitionnement peut être résumé dans la figure 3.4. Ce partitionnement a été effectué cinq fois afin de s’assurer que les caractéristiques sélectionnées ne sont pas trop sensibles au partitionnement des données.

0.3.8 Scénarios de régression

Comme nous nous intéressons au monitorage de la gravité de la MA, une tâche de régression est nécessaire. Plus précisément, nous souhaitons estimer le score MMSE d’un patient atteint de la MA. Motivé par les connaissances de [18], trois régresseurs supervisés sont explorés [51] : La régression par machine à vecteurs de support d’appui avec des noyaux linéaires (SVM-Linear) et gaussiens (SVM-RBF), la régression par forêt d’arbres décisionnels (RF) et la régression par la méthode des plus proches voisins. Tous les modèles ont été implémentés avec la bibliothèque open-source scikit-learn écrite en python [79].

0.3.9 Mesure d’évaluation

Pour évaluer la performance de chaque modèle, nous utilisons comme mesure l’erreur quadratique moyenne (EQM) entre la prédiction et le score réel du MMSE. Comme l’ensemble de données disponible n’est pas très important, nous nous appuyons sur une validation croisée répétée de k fois comme dispositif de test. De cette façon, l’ensemble de données disponible pour le train/le test (c’est-à-dire 75 % de l’ensemble de données original) est divisé de manière aléatoire en k plis (par répétition). La formation est effectuée avec k-1 plis et ensuite testée sur le pli tendu. Cette opération est répétée jusqu’à ce que tous les plis aient été utilisés comme plis maintenus. Avec la validation croisée répétée, cette procédure est répétée N fois, chaque fois avec une partition différente. Au final, un total de NK RMSE est produit. Ici, $N = 50$ et $K = 5$ sont utilisés. La moyenne et l’écart-type de ces valeurs NK sont utilisés pour évaluer les performances du système. Idéalement, nous nous intéressons aux modèles qui obtiennent une valeur efficace moyenne faible, ainsi qu’un faible écart-type, ce qui suggère que le modèle est insensible au partitionnement des données.

0.4 Résultats et discussions

0.4.1 Analyse statistiques des caractéristiques

Dans un premier temps, nous avons voulu évaluer la relation entre les résultats du MMSE et les caractéristiques de l'EEG et de l'IRM extraites ici. À cette fin, la corrélation de Spearman a été utilisée et seules les caractéristiques ayant une valeur p inférieure à 0,05 ont été retenues pour l'analyse. Les tableaux 4.1 et 4.2 indiquent les coefficients de corrélation de Spearman pour les caractéristiques EEG et IRM, respectivement. Comme on peut le voir, 39 caractéristiques EEG ont validées le critère de signification, et 29 pour les caractéristiques IRM. Les valeurs absolues de corrélation vont de 0,21 à 0,31 pour les caractéristiques EEG et de 0,21 à 0,37 pour l'IRM, ce qui suggère qu'individuellement, l'IRM peut fournir des détails plus pertinents pour la progression de la MA. Il est clair que les caractéristiques de modulation d'amplitude de l'EEG ont montré la plus grande corrélation, tandis que les caractéristiques de parcellisation du cortex et de la substance blanche se sont avérées les meilleures pour l'IRM. Sur les dix caractéristiques principales de l'EEG, trois correspondaient aux nouvelles caractéristiques de la parcellation spectrale de modulation, ce qui suggère qu'elles peuvent effectivement être utiles pour la régression.

Puisque des travaux antérieurs ont également suggéré une corrélation entre les caractéristiques de l'EEG et de l'IRM [?], nous explorons ici la relation entre les modalités des caractéristiques. La figure 4.1 représente une carte thermique selon le coefficient de corrélation de Spearman pour toutes les caractéristiques des modalités d'EEG et d'IRM. Les caractéristiques de l'EEG sont représentées par le grand rectangle noir sous l'EEG, et il montre le coefficient de corrélation entre les variables de l'EEG. Par conséquent, le plus petit rectangle noir au-dessus de l'IRM apporte la corrélation associée par les caractéristiques extraites des scans IRM. Les zones qui n'appartiennent à aucun des rectangles noirs représentent la corrélation entre les caractéristiques de l'EEG et de l'IRM. Comme on peut le voir, il existe une certaine relation entre les modalités, mais les corrélations entre les caractéristiques au sein d'une même modalité sont sensiblement plus élevées, ce qui suggère que les groupes de caractéristiques 1, 2 et 4 peuvent conduire à de meilleurs résultats.

0.4.2 Prévision de la gravité de l'AD

Ici, nous explorons les avantages de la prévision de la gravité de la MA à l'aide de systèmes unimodaux et multimodaux. Les résultats expérimentaux pour les quatre groupes de caractéristiques sont présentés dans le tableau 4.3, où le score moyen de l'EMR suivi de son écart-type est indiqué sur l'ensemble des essais de validation croisée et dont la moyenne est calculée sur les cinq séries de tests. En outre, les figures 4.2, 4.4, et 4.3 présentent les distributions des scores moyens de l'EMR en fonction des différents algorithmes de régression, des différents groupes de caractéristiques, et des différents algorithmes de sélection des caractéristiques, respectivement.

En outre, dans le but de déterminer si les différences de performance observées sont statistiquement significatives, un test statistique d'ANOVA à sens unique a été effectué pour différentes populations de l'EMR pour chacune des divisions, comme le montrent les figures 4.2, 4.4, et 4.3. Tous les résultats étaient significatifs à $p < 0,05$.

Comme on peut le voir dans le tableau et la figure 4.2, les résultats obtenus suggèrent que le SVM-RBF et le classifieur à arbres décisionnels obtiennent des erreurs de prédiction plus faibles par rapport à la SVM avec un noyau linéaire ou un régresseur k-NN. Le classificateur SVM-RBF, à son tour, a obtenu une variabilité RMSE plus faible, ce qui suggère que le modèle est moins sensible au partitionnement des données, un facteur important pour la capacité de généralisation du système. En outre, les résultats du tableau et de la figure 4.3 suggèrent que les algorithmes de sélection des caractéristiques ne jouent pas un rôle crucial (plus de détails sur les caractéristiques principales dans la section suivante) sur les performances finales. En ce qui concerne les groupes de caractéristiques (voir la figure 4.4), entre les systèmes unimodaux (Groupe 1 et Groupe 2), les caractéristiques IRM ont obtenu des valeurs RMSE légèrement inférieures par rapport aux caractéristiques EEG, corroborant ainsi les récentes découvertes [39]. Dans l'ensemble, pour la plupart des cas testés, les systèmes multimodaux, en particulier ceux formés sur les caractéristiques du groupe 4, ont montré de légères améliorations de la performance par rapport aux systèmes unimodaux. En particulier, une variabilité plus faible de la RMSE a été observée, ce qui suggère que les systèmes multimodaux sont moins sensibles au partitionnement des données. Bien qu'un régresseur à arbres décisionnels avec les caractéristiques du groupe 4 ait donné la RMSE moyenne la plus faible (1,827), la variabilité était assez importante. Pour des applications pratiques, cependant, une combinaison de moyenne RMSE et de variabilité plus faible est souhaitable. En tant que telle, la méthode SVM-RBF, combinée avec

les caractéristiques du groupe 3 ou du groupe 4, s'est avérée être la solution la plus prometteuse en terme de performances.

0.4.3 Caractéristiques pertinentes

Afin de mieux comprendre l'impact des différentes modalités des caractéristiques sur la performance globale du monitorage de la gravité, une analyse approfondie est effectuée concernant les caractéristiques qui ont été systématiquement sélectionnées au cours des cinq tests de configuration. À cette fin, les tableaux 4.4, 4.5, et 4.6 présentent la liste des caractéristiques qui ont été sélectionnées plus d'une fois par chacune des techniques de sélection des caractéristiques pour les groupes de caractéristiques 1-3, respectivement.

Les tableaux montrent que le MRMR s'est avéré moins sensible au partitionnement des données par rapport aux autres techniques de sélection des caractéristiques, car il a pu sélectionner des caractéristiques similaires de manière cohérente au cours des cinq passages. Cela a été le cas pour tous les groupes de caractéristiques testés. Il s'agit d'un aspect important qui doit également être pris en compte lors de l'évaluation des performances globales du système [62]. Ceci, combiné à l'amélioration des performances obtenue avec les régresseurs SVM-RBF, suggère que le système le plus fiable pour la détection de la gravité de la MA serait un système basé sur la sélection de caractéristiques MRMR, et le régresseur SVM-RBF et sur un ensemble de caractéristiques multimodales (groupe 4).

D'après les analyses du tableau 4.4, les caractéristiques *kurtosis – R2oR3 – P3 – P4*, *skewness – beta – mdelta – F4*, *cv – theta – mdelta – P3*, et *cv – R2oR3 – P3 – P4* se sont avérées être les caractéristiques les plus souvent sélectionnées, soulignant ainsi l'importance des caractéristiques spectrales de modulation non seulement pour le diagnostic, mais aussi pour le monitorage de la gravité de la MA. L'électrode virtuelle *P3 – P4* s'est également révélée importante, corroborant ainsi le déficit inter-hémisphérique couramment signalé dans le cas de la MA [66]. Plus important encore, parmi les nouvelles caractéristiques de patchs spectraux de modulation proposées, elles représentaient la moitié des 20 premières caractéristiques d'EEG sélectionnées par le MRMR, soulignant ainsi leur importance pour les systèmes de monitorage unimodaux basés sur l'EEG.

Dans le tableau 4.5, un plus grand nombre de caractéristiques dérivées de l'IRM ont été sélectionnées de manière cohérente, par rapport au groupe 1. Les zones mesurées à partir du pôle frontaux gauche et du temporaux supérieur ont été sélectionnées dans toutes les passes par la technique MRMR, corroborant ainsi l'importance des lobes frontaux et temporaux pour le diagnostic de la MA [61]. En outre, si la MA est principalement associée à ses effets sur la matière grise corticale, des dommages à la matière blanche ont également été signalés avec la MA. Les auteurs de [13], par exemple, ont trouvé des différences dans les scans des patients atteints de la MA par rapport aux sujets sains dans la substance blanche des lobes frontaux, temporaux et pariétaux. Dans ce cas, le MRMR a sélectionné le volume du masque de la substance blanche, le volume de la substance blanche cérébrale et a estimé le volume intracrânien total dans quatre des cinq passages. Les mesures du côté droit du pôle frontal, du supramarginal, du temporal supérieur, du pariétal supérieur, du temporal transversal, du volume de la substance blanche cérébrale, de la substance blanche non segmentée et de l'insula ont également été sélectionnées dans quatre des cinq séries, en plus des mesures du côté gauche du volume de la substance blanche cérébrale et de la zone entorhinale. En outre, les zones des côtés gauche et droit de l'occipital latéral et les mesures de la substance blanche du côté droit de l'occipital latéral et des fusiformes ont été sélectionnées au moins quatre fois à partir de la technique de sélection des caractéristiques de corrélation de Pearson et Spearman. Ce sont toutes des zones et des caractéristiques qui ont été associés à la MA (par exemple, [39, 67, 12]).

Enfin, le tableau 4.6 présente les caractéristiques du groupe 3 sélectionnées de manière cohérente. Sur les 29 caractéristiques qui ont été sélectionnées dans plus d'un cycle, seules 6 proviennent de l'IRM. Les caractéristiques les plus importantes sont également apparues précédemment lors des tests unimodaux, ce qui suggère leur importance pour la tâche à accomplir.

Les tableaux 4.7 et 4.8 présentent le résumé global des caractéristiques sélectionnées de manière cohérente en fonction de la catégorie de caractéristiques, des bandes de fréquences, des positions des électrodes et des hémisphères cérébraux. Parmi les caractéristiques de l'EEG, les caractéristiques d'asymétrie se distinguent dans les systèmes unimodaux, tandis que les caractéristiques d'écart-type se distinguent pour les systèmes multimodaux. Ces résultats suggèrent que la dynamique temporelle des caractéristiques joue un rôle important dans le monitorage de la gravité, ce qui est corroboré par l'importance observée avec les caractéristiques de taux de changement de modulation. En fait, tant pour les systèmes unimodaux que multimodaux, les caractéristiques de modulation

d'amplitude se sont révélées extrêmement importantes, en particulier les caractéristiques liées à la dynamique de la sous-bande de fréquence bêta, corroborant ainsi les recherches précédentes [56]. Les régions de modulation $R2/R3$, à leur tour, se sont révélées extrêmement importantes pour les systèmes unimodaux. En ce qui concerne la position des électrodes, les canaux bipolaires virtuels se sont révélés importants pour les systèmes unimodaux, ce qui est conforme aux recherches précédentes qui ont montré une déconnexion inter-hémisphérique avec la DA [57]. En ce qui concerne les caractéristiques de l'IRM, les caractéristiques dérivées des groupes APARC et WM ont été plus souvent sélectionnées plus d'une fois que les caractéristiques du groupe ASEG. De plus, les caractéristiques du côté droit du cerveau apparaissaient plus souvent dans les systèmes unimodaux, corroborant ainsi les résultats de la littérature sur la dominance droite et l'asymétrie de latéralité (par exemple, [70, 40]).

Comme dernière expérience, nous avons appris des classificateurs en utilisant uniquement les caractéristiques sélectionnées de manière cohérente à partir de l'algorithme du MRMR, comme indiqué dans les tableaux 4.4, 4.5 et 4.6 pour l'EEG uniquement, l'IRM uniquement et l'EEG-MRI, respectivement. Le tableau 4.9 présente les résultats obtenus à partir de chacun des quatre régresseurs dans un schéma de test de validation croisée. Comme précédemment, le SVM-RBF et le classificateur d'arbres décisionnels ont obtenu les meilleures performances, ce qui souligne encore l'importance des modèles non linéaires. Le SMV-RBF a obtenu la variabilité la plus faible parmi les essais de validation croisée. En outre, pour les systèmes unimodaux, ceux basés sur l'IRM ont, là encore, surpassé ceux basés sur l'EEG. Néanmoins, les systèmes multimodaux se sont avérés toujours plus performants que chaque système unimodal pris individuellement. Le meilleur système global a été obtenu avec un système multimodal et un régresseur d'arbres décisionnels, atteignant une RMSE moyenne de 1,682.

0.5 Conclusions

La maladie d'Alzheimer touche des millions de personnes dans le monde entier, et il est probable que ce nombre augmentera encore dans un avenir proche. Malheureusement, la maladie est encore incurable et irréversible, c'est pourquoi les méthodes de diagnostic précoce et de suivi automatique de l'évolution de la gravité de la maladie ont gagné en importance. Cette thèse a exploré le potentiel

de la construction d'une méthode multimodale automatisée pour mesurer les niveaux de gravité de la maladie chez les patients ayant reçu un diagnostic de maladie d'Alzheimer légère.

Ici, pour la première fois, de nouvelles fonctionnalités de diagnostic basées sur l'EEG ont été explorées pour le suivi de la gravité de la maladie, ainsi que leur fusion avec des biomarqueurs IRM. Quatre algorithmes d'apprentissage machine différents ont été étudiés, ainsi que trois méthodes permettant de passer au crible les caractéristiques les plus pertinentes pour la tâche à accomplir. Les résultats expérimentaux ont montré que si les nouvelles caractéristiques de modulation spectrale de patchs EEG sont plus pertinentes que les caractéristiques de puissance spectrale conventionnelles pour la surveillance de la gravité, les systèmes formés aux caractéristiques de l'IRM ont eu tendance à surpasser les caractéristiques de l'EEG. Dans l'ensemble, les systèmes EEG-MRI multimodaux ont montré une meilleure précision et une variabilité plus faible lors des validations croisées, ce qui suggère leur importance pour les applications pratiques. Parmi les différents algorithmes de sélection des caractéristiques, l'algorithme MRMR a montré une meilleure stabilité lors des différents tests de configuration et s'est distingué des autres candidats basés sur la corrélation. Parmi les algorithmes d'apprentissage machine, les régresseurs non linéaires ont montré une meilleure précision que les régresseurs linéaires.

Les résultats présentés dans cette thèse ont démontré le potentiel du développement d'un système automatique capable de prédire le score MMSE d'un patient atteint de la maladie d'Alzheimer au stade léger, en utilisant des biomarqueurs extraits de l'EEG et de l'IRM. En outre, l'étude pourrait stimuler et encourager la recherche et le développement d'un tel système. Ainsi, les cliniciens, les patients et les chercheurs pourraient avoir accès à un système automatique d'assistance pour mesurer l'évaluation cognitive.

0.5.1 Limites de l'étude

Les résultats présentés dans cette étude ont été obtenus sur un échantillon limité de 89 participants. Bien que des validations croisées répétées et des essais multiples aient permis d'explorer les biais de partitionnement des données, la capacité de généralisation des modèles pour les patients invisibles est encore inconnue. En outre, à la connaissance des auteurs, il n'existe aucun ensemble de données en libre accès qui inclut les signaux EEG et les scanners IRM. Les données disponibles présentent une répartition très étroite des scores MMSE, comme le montre la figure 3.4. Bien que

cela permette de construire des modèles pour les patients présentant des niveaux légers de la maladie d’Alzheimer, on ignore encore comment les modèles se comporteraient à mesure que la maladie évolue vers des stades modérés et graves. C’est pourquoi une analyse plus approfondie est encore nécessaire. En outre, les performances des modèles ont montré une amélioration lorsque des caractéristiques présentant une stabilité dans les séries d’expériences de sélection de caractéristiques ont été utilisées. À cet égard, des techniques plus récentes de sélection de caractéristiques devraient être explorées dans la recherche et l’identification de caractéristiques sélectionnées plus stables, en particulier en utilisant des méthodes d’enveloppe dans lesquelles la sélection de caractéristiques et le régresseur sont étroitement couplés. Enfin, grâce aux progrès réalisés dans l’analyse des réseaux neuronaux profonds de [85] et de l’IRM [58], des architectures multimodales profondes devraient être explorées une fois que des données supplémentaires seront mises à la disposition de la communauté des chercheurs.

0.5.2 Travaux futurs

Quant aux futures orientations de la recherche, les travaux pourraient se concentrer sur l’amélioration des performances des modèles prédictifs automatiques par un meilleur réglage des hyperparamètres, l’utilisation et l’exploration de différentes stratégies de fusion, et le développement de nouvelles fonctionnalités. Plus important encore, l’évaluation des résultats obtenus à l’aide d’un ensemble de données comprenant des patients présentant une propagation du MMSE beaucoup plus importante serait nécessaire pour explorer si les modèles se généralisent bien aux stades plus avancés de la maladie. En outre, avec la disponibilité d’ensembles de données plus importants, des architectures de réseaux neuronaux profonds de pointe pourraient également être explorées.

Chapter 1

Introduction

Doctor Alois Alzheimer investigated, at the beginning of the twentieth century at the Frankfurt Psychiatric Hospital, the symptomatic progression of a female patient registered as Auguste D., who presented at the time sleep disorders, disturbances of memory, aggressiveness, crying, and progressive confusion. After Auguste D.'s death, Dr. Alzheimer was able to analyze her brain tissue and found morphological differences, which he then connected to the clinical symptoms observed. This marks the discovery and the beginning of what would later be called Alzheimer's disease [54].

Nowadays, researchers understand Alzheimer's disease (AD) as a progressive degenerative brain disorder, which is considered to be the most common cause of dementia. The characterization of dementia is a conjunction of symptoms such as functional decline in memory, language, problem-solving, and other cognitive skills that can directly impact the individual's quality of life. The symptoms appear due to the death or damage of brain nerve cells (neurons), which could start twenty years or more before the evidence of the symptoms [3]. AD is a neurodegenerative disorder that is likely to reach people over 65 years of age and is still, unfortunately, incurable [2].

Studies showed that in 2016 a population of 26.6 million subjects were living with AD worldwide. Additionally, models predict that by 2050 this number will be multiplied by four, meaning that 1 in every 85 people will have developed AD. Substantially, as the life expectancy of the population increases, the number of aged people also increases. One way to prevent the world from facing the ongoing global epidemic of AD could be the advances in preventive planning and therapies to slow

down AD progression since the estimation is that 43% of the cases in 2050 will require high levels of care [14].

Individuals suffering from AD could develop a variety of symptoms, and depending on the stage of the disease, their intensity can vary. Some symptoms of the disease usually include fast intellectual and behavioural decline, which are usually associated with short-term memory difficulties. As detailed in [2], symptoms typically dispose four primary modalities. Physiological symptoms include changes in brain tissue due to the presence of plaques and tangles, cerebral hypoperfusion, and atrophy in the brain cortical and hippocampal areas. Depression, changes in mood, and apathy, in turn, are some examples of the psychological symptoms faced by AD patients. Cognition symptoms consist of memory impairments, obstacles in solving daily problems, and communication problems. Finally, as a consequence of the previously mentioned symptoms, the individual living with the disease may present incoherent behaviour.

Researchers are making efforts to identify risk factors that can be associated with the disorder [35]. Some of the risk factors linked to the development of AD include:

1. Age: Ageing is considered one of the most significant risk factors for the development of AD. Individuals that develop the disease are predominantly 65 years old or older. Additionally, the prevalence of the disease increases as the individual ages [53].
2. Family history: Indicated as the second most severe risk factor for AD by [96], family history follows advance in aging in the list. Individuals with an immediate relative with AD have higher probabilities of developing the disease than individuals who do not [3].
3. Sex: While men and women seem to develop AD in the same proportion until a later age, the number of women developing AD later in life is higher than the number of men. The more significant susceptibility of developing the disease by women at a later age could be associated with females living longer than males on average [83].
4. Education: Studies have shown that people with more years of education have less possibility of developing AD and other dementias than individuals who have had fewer years of schooling [88].

5. Others risks: Cerebrovascular diseases, hypertension, type 2 diabetes, obesity, dyslipidemia, smoking, stress, depression, inadequate sleep, lack of physical activity, and diet are among others in the list of risk factors of developing AD [93].

1.1 Phases of Alzheimer's disease

The literature reports that individuals with AD go through three phases: preclinical, mild cognitive impairment, and dementia due to AD [18, 2]. The disease begins way before any of the symptoms are apparent in the first stage denominated by preclinical AD. By virtue of a lack of symptoms, it is hard to detect the disease at this stage. However, changes in the brain, blood and cerebrospinal fluid (CSF) may already be happening.

The second stage of the disease is called mild cognitive impairment (MCI). In this phase, problems related to thinking capability, as memory lapses, difficulty in choosing the right vocabulary, and geographical location of familiar places become apparent. Even though the symptoms are noticeable for the patients themselves and/or people around them, the independence of the individuals in daily life activities is not affected in this stage. Not everyone that is diagnosed with MCI will develop dementia due to AD later on, however. The reason is still unknown as to why some people develop dementia while others do not.

In the last stage of the disease, patients develop dementia due to AD. At this instant, memory, thinking, and behavioural impairments are very noticeable, considerably affecting the patient's daily life activities. Two different sub-stages may derive from dementia due to AD, namely, moderate or middle-stage AD and severe or late-stage AD. The former refers to the most extended phase of the disease, where the execution of daily life tasks becomes an issue for the individuals, which leads to frustration and anger. The latter is the final stage where patients' cognitive and memory skills are damaged, thus requiring 24-hour personal assistance, since the patients' movements and physical abilities are also affected.

1.2 Diagnosis and treatment of Alzheimer's disease

Currently, a precise diagnosis of AD is achievable with a postmortem structural examination of the brain. Traditionally, diagnosis has relied mostly on cognitive assessment using paper and pencil tests, such as the Montreal Cognitive Assessment (MoCA) [76] and/or the Mini-Mental State Examination (MMSE) [11]. With the MMSE, for example, a professional clinician asks the patient several questions about daily mental competencies to evaluate their cognitive state. The participant is then given a score based on their answers, with a maximum of 30 points [45]. Scores of 20-24 are indicative of mild dementia, whereas scores of 13-20 suggest moderate dementia, and those less than 12 are indicative of severe dementia.

These tests typically detect cognitive impairment once damage to the brain is substantial. Research is therefore on the rise to find biomarkers of the disease before symptoms arise. It is thought that disease modifying therapeutic interventions should be most effective when administered in the early stages of the disease, before neuronal loss occurs [55], thus early detection is crucial. Over the last two decades, a number of biomarkers have been suggested, including those based on structural neuroimaging (e.g., hippocampal volume from magnetic resonance imaging) [1, 48], cerebral spinal fluids (e.g., tau and beta-amyloid levels) [16], blood and urine samples [74], electroencephalography [18], and more recently, genetic risk profiling [99].

As Alzheimer's disease is still an incurable disorder, treatment is usually multifaceted and based on the management of the symptoms and the slowing of the clinical decline. For a favourable treatment, the first step is the awareness of the symptoms followed by a proper diagnosis. At this point, the patient and caregivers need to be educated about the disease, the expected course, treatment options, and expectations regarding disease development. Non-pharmacological approaches can be suggested based on patient and caregivers' needs, priorities, and resources, to alleviate symptoms such as agitation, apathy, delusions, disorientation, resistance to care, among others. Subsequently, the establishment of a suitable pharmacological treatment may be necessary. In the long term, the care plans need adjustment and modification according to the changes and conditions of the patient as well as of the caregivers [5].

1.3 Biomarkers for Alzheimer's disease

Biomarkers refer to signals collected from a patient for medical diagnostic purpose. Usually used as indicators of biological mechanisms, biomarkers need to be measured accurately and in a reproducible manner. Biomarkers are commonly used in areas such as drug development, disease diagnostics, and even to better understand healthy physiology. Once the relationship between a biomarker of biological mechanisms and a specific clinical process is defined, the relationship can be used to advance disease treatment, for diagnostic, as well as to help monitor disease progression [95]. For AD, as mentioned before, a large group of biomarkers have been developed, ranging from brain imaging, electrophysiology, to cerebrospinal fluid analyses [50]. The list presented below includes some of the common biomarkers used for AD.

1.3.1 Cerebrospinal fluid (CSF)

Due to the presence of CSF flowing around the human brain and spinal cord, researchers have investigated if the fluid could demonstrate any changes occurring in the brain as a result of the development of AD. According to [52], there are four well established CSF biomarkers, namely, A β 40, A β 42, total tau, and phosphorylated tau that have contributed positively to the diagnostic procedure of AD. The main drawback of CSF biomarkers lies on its invasiveness to collect the samples, as a lumbar puncture is needed [18].

1.3.2 Positron emissions tomography (PET) scans

PET scans offer three-dimensional visual information about the glucose metabolism in the brain. Since patients with AD have shown a degradation in their glucose metabolism, studies have explored PET to identify valuable functional changes occurring in the brain, not only for diagnostics, but also for AD severity measurement [77]. The main drawback is the need for radioactivity exposure [2].

1.3.3 Magnetic resonance imaging (MRI)

Structural MRI produces brain imaging as a result of applying a powerful magnetic field to the brain area. It generates images of different brain tissues at a set moment in time, as tissues generate a different response to the field applied. This technique is able to detect anomalies encountered in the brain's anatomy due to the atrophies caused by the disease in some areas of the brain. A second utilization of MRI consists in measuring the concentration of oxygen in a determined area of the brain when the individual performs a specific task. This technique, termed functional MRI (fMRI), detects abnormalities in the functionality of the AD patient's brain [2].

1.3.4 Electroencephalography (EEG)

EEG is responsible for recording the electrical field of the scalp produced by the interaction of neurons. Several studies over the last couple of decades have shown varying neurodegenerative effects of AD on the EEG signals. Given the lower cost and portability of the technology, amongst other advantages, this has made EEG into a promising instrument to measure biomarkers to support AD diagnosis and progression monitoring [18].

1.3.5 Magnetoencephalography (MEG)

MEG records brain activity from the same source as measured by EEG, with the difference that it measures the magnetic field produced by the neurons' electric current movements, as opposed to the electrical fields per se. The number of MEG studies applied to AD have been significantly lower than those relying on EEG, likely due to cost and lengthy setup times required and with obtained results in line with that achieved by a simpler EEG setup [2].

1.4 Alzheimer's disease severity

Upon receiving an AD diagnosis, patients and their families start to raise questions regarding the progression of cognitive decline. Unfortunately, the rate of progression for Alzheimer's disease varies widely, and on average, people with the disorder can live between three and 20+ years after

diagnosis, thus leaving the clinicians with a difficult task of monitoring disease progression [33]. Moreover, as more therapies, drugs, and interventions appear, monitoring the progression and/or slowing of the disease will become crucial to gauge the benefits of the different interventions [101]. This is particularly challenging with early detection, as symptoms may not be as pronounced.

While several studies have shown the benefits of biomarkers for early AD detection of AD, fewer have explored their usefulness for monitoring disease progression. In [55], for example, several biomarkers from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database were explored and hippocampal volume, tau and beta-amyloid levels were reported to be promising. In [36], in turn, hippocampal atrophy and cortical thickness of a selection of temporo-parietal regions were shown to be useful predictors of mild cognitive impairment progressing to AD. More recently, in [18], it was shown that only a few studies have explored the effects of disease progression on EEG. In particular, earlier work showed that changes in specific frequency subbands could be seen with disease progression (e.g., [64, 65]). More recently, spectro-temporal energy patches were shown to provide improved accuracy (relative to conventional subband powers) when predicting disease severity levels [19].

1.5 Thesis objective

As Alzheimer’s disease manifests itself across different facets, it is expected that a multimodal biomarker system would be needed for accurate Alzheimer’s disease progression monitoring [52]. In fact, some studies suggest that a single biomarker may not provide sufficient information to determine disease severity accurately [50]. While multimodal biomarkers have been explored for AD and MCI detection [102], most have relied on structural neuroimaging tools [37] and only a handful have explored the fusion of structural neuroimaging with electro-neurophysiology [18]. Moreover, limited work has been explored for multimodal systems for AD severity detection [73].

More specifically, this thesis explores the development of a multimodal correlate of AD severity level for patients at early stages of the disease (MMSE scores between 21-26). We explore the fusion of several structural, cortical thickness and white matter features extracted from MRI, as well as conventional EEG features alongside ones recently developed, such as the spectro-temporal energy

patch features [19]. Experiments with 89 patients show the usefulness of multimodal AD severity level models for patients at early stages of AD.

1.6 Publications

1. Exploring Predictive Models of Alzheimer's Disease Severity based on Resting State EEG and MRI Features. B. Jesus, R. Cassani, W. McGeown, T. Falk, abstract, CCD'2019 [100].
2. Publication in preparation to be submitted to the special issue Brain-Computer Interface Applications for Improving the Quality of Elderly Living of the Computational Intelligence and Neuroscience journal .

1.7 Thesis organization

This thesis is comprised of four additional chapters, as detailed next.

- Chapter 2 presents a summary of typical EEG and MRI biomarkers used in AD research. Moreover, the chapter presents a small overview of the components involved in the disease severity monitoring system.
- Chapter 3 describes the methods and materials used throughout the implementation of the study. We start with a description of the data used, the pre-processing analyzes applied, the features extracted, the feature selection approaches used, and finally, the regression models implemented in the predictive systems.
- Chapter 4 presents the results obtained with the different uni- and multimodal models, as well as an analytical discussion based on the obtained results.
- Chapter 5 presents the thesis conclusions, including the limitations and difficulties the study faced, as well as potential future research directions.

Chapter 2

Biomarkers and AD severity monitoring system

2.1 Magnetic resonance imaging

Driven from the science called by nuclear magnetic resonance (NMR), three scientists, Raymond Damadian, Paul Lauterbur, and Sir Peter Mansfield, involved themselves in studies connected with the formulation of a mechanism to create images of body structures which were more detailed than those produced at the time, leading to the so-called MRI. The technique changed considerably the medical imaging world and has continuously been upgraded, improved, and applied to numerous contemporary clinical problems [41].

In summary, the MRI technique aligns the nuclear spins of the hydrogen atoms present in the human body by applying a powerful, dominant external magnetic field. The perturbation of the nuclear spins' state of equilibrium is caused by the injection of energy in form of radio-frequency (RF) pulses. Thus, the injection of RF energy will distort the longitudinal magnetization (the magnetization occurring in the same axis of the dominant magnetic field) and provokes the transverse magnetization (perceptible magnetization occurring in the perpendicular axis of the dominant magnetic field). Supplemental gradient magnetic fields are used to generate secondary magnetic fields in order to distort the primary dominant magnetic field so the system is capable of generating images in the X, Y and Z axes. After the pulse, the spins will return to the state

prior to when the perturbation was applied, releasing energy. Then, antennas, coupled to the MRI structure, sense the released energy. Finally, interpretation of the desired information might derive from the duration of the spins' realignment (relaxation), the proportion of the released energy, and the differences in the tissues' magnetic properties. Information can be produced by two different relaxation times: T1, which is associated with the time spent for the tissue to recover its original longitudinal magnetization, and T2, which is associated with the time consumed for the dissipation of the transverse magnetization [89].

MRI is broadly used in the clinical scenario to assist physicians in diagnosing diseases, identifying injuries, and other applications. Examples of clinical problems involving the utilization of MRI include diagnosis of mild traumatic brain injuries [91], brain tumor investigations [8], assessment of cardiac structures [29], clinical ophthalmological uses [97], and diagnosis and treatment of cancer [98]. MRI is also pointed as a potential biomarker to determine symptomatic effects and treatment effects by measuring the rate of atrophy of the brain regions [61].

2.2 MRI and Alzheimer's disease

Structural brain MRI has been extensively studied as a neuroimaging tool to help diagnose AD [2]. Research progress has enabled the development of automated methods that can perform subcortical segmentation of brain regions and determine measures of cortical thickness and volume, critical measures that help detect AD biomarkers [61]. MRI has also shown to be a potential tool for disease progression monitoring via volumetric measurements [52].

Several studies have shown the ability of structural MRI in assessing structural atrophy in some distinctive areas of the brain. Reduced volume of the hippocampus and entorhinal cortex are the most persistent markers encountered in MRI scans of patients with AD. Hippocampal atrophy, however, is not distinct to AD and has also been associated with other pathologies, such as schizophrenia and depression [52], thus the need for multimodal systems.

The literature also reports decreased gray matter volume in the parahippocampal gyrus, in the amygdala, in the posterior association cortex, and the subcortical nuclei as markers of AD [52]. Measures of volume and cortical thickness of the precuneus and posterior cingulate cortex (PCC), parietal lobe, temporal, and frontal lobe have also been related to the AD neuropathology [61].

Many researchers have been committing their efforts in the development of an automatic diagnostic system for Alzheimer's disease since it would be a significant assistant to clinicians and patients. Several such systems have relied on magnetic resonance imaging, such as [81, 7].

2.3 Electroencephalography

In the early 90s, after the German psychiatrist Hans Berger identified that low-level sub-audio-frequency electrical activity produced by the human brain could be registered, electroencephalography measurements were established. Since then, EEG has become a handy tool in the exploration and discovery of brain functionality. The technique has turned into an attraction to scientists of diverse fields, ranging from medicine to engineering [92].

EEG signals record the electrical activities caused by the generation of secondary electrical fields produced by currents that flow at synaptic excitation from thousands of neurons, with similar spatial orientation, in the cerebral cortex near the scalp. Electrodes positioned on the scalp of a subject can record potentials caused only by a large population of active neurons, since it is necessary to overcome skull attenuation and brain tissue [90].

Due to the EEG property of recording changes in the brain operability, the technique has been employed in studies to diagnose many neurological disorders and abnormalities in the human brain. Examples of clinical problems involving the utilization of EEG include monitoring brain death activity [49], cognitive engagement [10], investigation of brain injuries [78], detection of epilepsy [60], assessment and testing of drugs [87], and investigation of other mental disorders [28, 63, 86].

2.4 EEG and Alzheimer's disease

EEG has shown to be a promising resource in the diagnosis and investigation of AD due to its advantages regarding portability, price, and non-invasiveness compared to the other techniques. Encouraged by its advantages, researchers have been studying in what way AD and its progression could affect EEG signals. A recent systematic review of the literature has identified four consistent effects of AD on resting-state EEG recordings [18], namely:

2.4.1 Slowing of EEG

A significant effect observable in EEG signals is associated with a higher power in low frequency bands, such as delta and theta, and lower power in high frequency bands, such as alpha and beta, thus suggesting a slowing of the EEG with AD [9]. Moreover, the shift seems to be proportional to the progression of the AD [18].

2.4.2 Reduced complexity

EEG signals collected from AD patients appear to be less complex than EEG acquired from the same age control subjects, which indicates that EEG signals from patients with AD are composed of more similar waveform arrangements while EEG signals from health subjects are more complex or “chaotic”. It is believed that this change in neural dynamics with AD is resultant from the increased loss of neurons, thus leading to decreased neuronal interaction and, consequently, EEG complexity [57].

2.4.3 Perturbation in EEG synchrony

Many studies have shown a decrease in EEG synchrony in patients with AD, suggesting a decrease in expected patterns in different brain regions and a loss of connectivity among certain cortical regions. The cause for this EEG synchrony perturbation is still not entirely understood [31].

2.4.4 Neuromodulatory deficit

Amplitude modulation analysis has been applied to measure quantitatively the neuromodulatory activity of EEG signals under resting states. Recent methods have shown neuromodulatory deficits and reduced cross-frequency dynamics with AD patients relative to age-matched controls [38, 46].

Similar to MRI-based approaches for the clinical diagnosis of AD, EEG could offer similar benefits, while providing a more portable, lower-cost diagnostic application. Examples of work that have been utilizing EEG in the development of an automatic diagnostic system include [21, 38].

2.5 Multimodal systems

A multimodal system corresponds to a diagnostic/severity level monitoring system that operates based on the combination of two or more user input modalities to produce the desired outcome. In [102], it is emphasized that the majority of AD diagnostic systems reported still rely on one individual modality. It is argued that biomarkers derived from different modalities can provide complementary information that may improve system robustness and accuracy.

To function as a multimodal system, the fusion of different input modalities needs to be performed. This can be achieved using three different methods, namely: data-, feature-, and decision-level fusion. Fusion is not trivial, as different modalities may have been collected under different sampling rates, with varying sensors and resolutions (e.g., 1-d time series, 2-d images, or 3-d images/tensors), and/or providing redundant information [34]. Next, a brief summary of the three fusion methods is given.

2.5.1 Data-level fusion

With this type of fusion, raw data from different sensors are combined. Generally, data-level fusion is applied when two modalities present a distinct similarity or relationship. This fusion method is susceptible to noise since the data is not always pre-processed beforehand. In contrast, data-level fusion integrates the complete set of information from the sensors, thus avoids loss of details during data processing. Data-level fusion is not very common in diagnostics, one the main reasons being the need for very accurate measurement synchrony between devices [24].

2.5.2 Feature-level fusion

Feature-level fusion, in turn, refers to the combination of modalities after the data has been processed independently to extract useful features/parameters. This processing makes the system more robust to noise compared to data-level fusion, but has the potential for information loss if discriminatory features are not extracted correctly [72]. Feature-level fusion is also challenging when features are extracted, for example, over different window durations (i.e., epoch sizes) or at different sampling rates.

2.5.3 Decision-level fusion

As the name suggests, decision-level fusion combines the decisions made from different systems, where one or more systems can be developed per modality and/or feature set. Decision aggregation can be done in many different ways, such as averaging of scores, taking majority vote, weighted-vote where weights can signal the quality of a given modality or the confidence of the system in its prediction, amongst others. Decision-level fusion is particularly useful when modalities are dissimilar (when data fusion is not possible) and/or when modalities have varying resolutions (when feature fusion is challenging). Of course, decision-level fusion requires some form of synchronization between systems in order to make a final decision (e.g., EEG-based systems can have decisions in the order of seconds, whereas fMRI based systems can be in the order of minutes) [72].

2.5.4 Hybrid fusion

Hybrid fusion, as the name suggests, can combine any of the three fusion methods described above into a final diagnostic system. For example, decision level fusion can be done between an MRI-based system and an EEG-based system, whereas the latter can itself be a system relying on feature-level fusion [4].

2.6 Automated prediction system overview

As mentioned previously, the goals of this thesis are two-fold: (1) build a multimodal system that combines EEG and MRI features for improved performance, and (2) to explore the use of a new EEG modulation spectrum patch feature for AD severity level measurement,. Figure 2.1 illustrates a high-level block diagram of an automated prediction solution . The image shows an unimodal system, but can be easily expanded to multiple modalities.

The first module consists of signal acquisition. This is the step in which raw data is collected from different types of sensors, stored, and made available for visualization and analysis. For this thesis, signal acquisition comprises the measurement of EEG signals during a resting state protocol and the acquisition of brain MRI scans. Next, since the quality of the data acquired is not always high and can be contaminated with artefacts, a pre-processing stage is applied. Here, the main

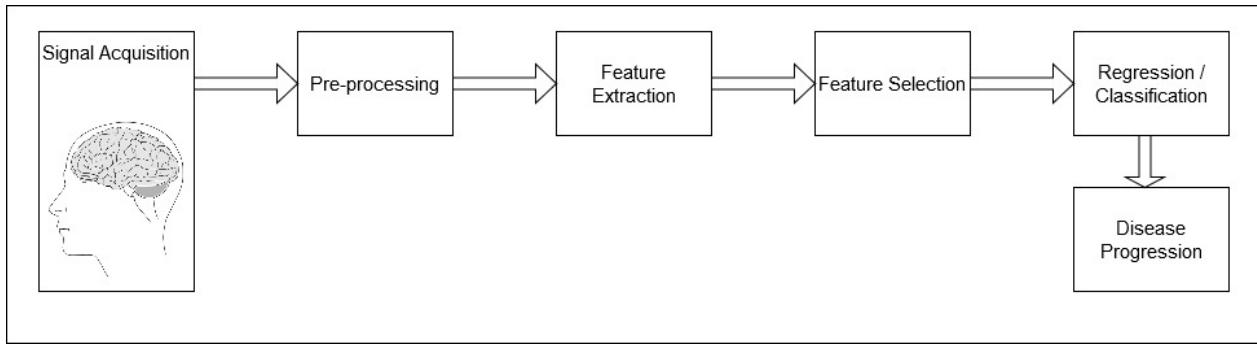


Figure 2.1 – Workflow of the prediction system.

objective is to enhance the quality of the raw acquired data via artefact removal and/or noise suppression algorithms.

Once the data has been enhanced and unwanted artefacts removed, the next stage corresponds to feature extraction, where important properties or information from the measured signals can be extracted. Typically, features are of much lower resolution than the original raw modalities (e.g., time-series or 2-d image) and convey information important for the downstream classification/regression task. Commonly, one does not know beforehand which features will provide the best discrimination between classes, thus numerous feature “candidates” are extracted. With this approach, however, many features may be correlated with each other, some may be more susceptible to noise than others, and some may provide no discrimination power, thus including them in the prediction system may make more harm than help. To this end, feature selection is employed such that the most relevant features for the task at hand are found and kept. By reducing the number of features, systems typically become less sensitive to database biases and are less prone to the so-called curse of dimensionality and to overfitting.

Once the most relevant features are found, the last stage of the prediction system comprises the classification or regression step. This is where models are trained to learn mappings between the top-selected features and the labels provided (assuming supervised learning is used). Classification is typically used in diagnostic systems, where discrete labels such as “control” or “AD” are used. Regression, in turn, aims to predict a continuous outcome variable, such as MMSE score; these are applicable to disease progression monitoring systems.

Chapter 3

Materials and methods

3.1 Participants

The data used herein was collected from a multi-site clinical trial exploring the use of event-related potential (ERP) biomarkers of Alzheimer's disease [26]. Data was originally collected from 103 subjects with probable mild AD and 101 healthy controls. The study was approved by the institutional review boards for each trial site and consent was obtained from each participant. The interested reader is referred to [26] for more details about inclusion and exclusion criteria, as well as all the tests and criteria performed for diagnostics.

Here, we are particularly interested in the target population data where ERP, resting-state EEG, and structural MRI scans were collected. After careful screening for sufficient scan quality for MRI segmentation/parcellation, a total of 89 participants were available for further analysis. Table 3.1 summarizes the demographics for these 89 participants. All patients presented an MMSE scores between 21 and 26.

Table 3.1 – Demographics for the patients used in this study

Subjects (female)	Age [years]	Education [years]	MMSE
89 (46)	75.8 ± 7.3	14.5 ± 3.3	23.3 ± 1.8

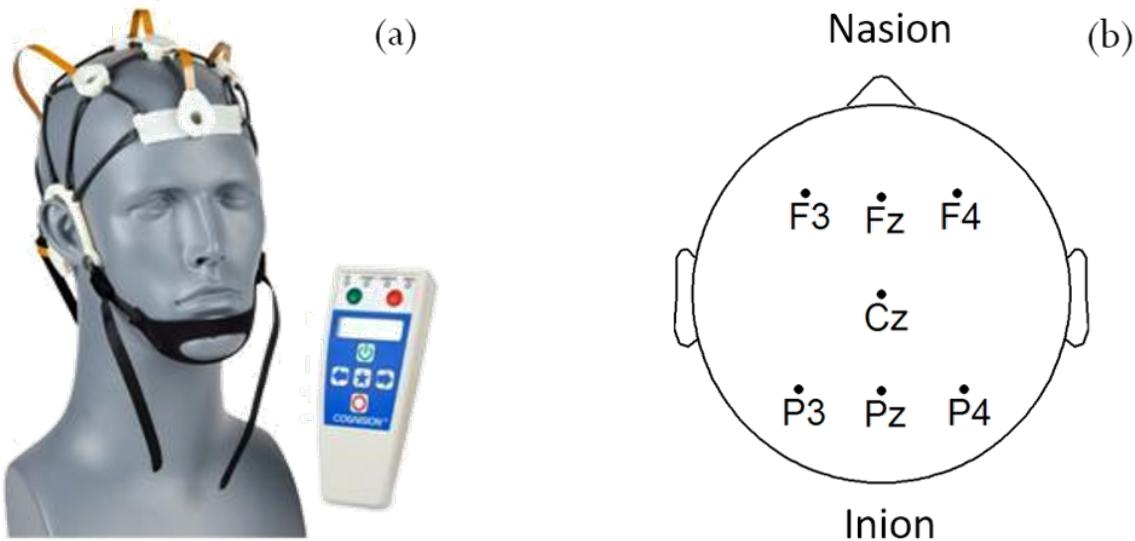


Figure 3.1 – (a) Cognision proprietary COGNISIONTM system. **(b)** EEG electrode placement for the AD dataset. Image taken from [21].

3.2 Data collection

EEG signals were acquired using the COGNISIONTM device [17], a seven-channel device operating at a sampling frequency of 125 Hz. Bi-auricular referential electrodes were also attached during the collection procedure. For each of the participants, we are interested in the EEG collected during a 3-min resting awake eyes-open period. Figure ?? displays the placement of the EEG electrodes located at the F3, Fz, F4, Cz, P3, Pz and P4 locations. Motivated from [38], two virtual inter-hemispheric bipolar signals are also used, namely: F3-F4 and P3-P4. MRI data with volumetry, in turn, was obtained from T1-weighted structural scans. The scans presented no movement or subsequent segmentation artifacts.

3.3 Pre-processing

Les signaux EEG étaient d’abord prétraités par un filtre passe-bande à réponse impulsionnelle finie (FIR) à phase zéro avec une bande passante de 0,5-45 Hz pour éliminer toute interférence avec les mesures du réseau électrique. En outre, sur la base des résultats présentés dans [22], une analyse des composantes indépendantes des ondes (wICA) a été appliquée pour éliminer les artefacts

oculaires et musculaires indésirables. Le lecteur intéressé est invité à consulter [71] pour plus de détails sur cet algorithme d'élimination des artefacts.

Les scans IRM pondérés T1, à leur tour, ont été traités par Freesurfer¹, une plateforme ouverte universellement utilisée pour traiter et analyser les scans IRM structurels du cerveau qui a été validée par les chercheurs IRM [42]. Les images capturées par le scanner IRM ont été prétraitées par la commande par défaut "reconall" fournie dans le pipeline Freesurfer pour mettre en œuvre la correction et la normalisation des mouvements.

3.4 Feature extraction

3.4.1 EEG features

In the context of Alzheimer's disease, four different features were computed from the acquired EEG signals based on insights from [21, 19]. These include: spectral power, magnitude squared coherence, amplitude modulation rate-of-change, and modulation frequency “patches” features. Features are extracted over 8-second epochs with 1-second shifts between consecutive epochs.

Spectral power

Spectral power features measure the power present in the EEG signal in specific frequency bands. For such, zero-phase FIR bandpass filters were used to decompose the EEG signal into different frequency bands of interest. Then, for each different subband time series $s(k)$ (with a total number of samples of m) the power is computed as:

$$\text{power} = \frac{1}{m} \sum_{k=1}^m s^2(k). \quad (3.1)$$

Next, normalization is performed by diving the subband power by the full-band EEG power. During our experiments, the spectral features were computed for the following frequency subbands: delta (0.1-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), low-alpha (8-10 Hz), high-alpha (10-12 Hz), beta (12-30 Hz), delta-to-beta (0.5-30 Hz), theta-to-beta (4-30 Hz) and low-gamma (30-45 Hz). A total

¹<https://surfer.nmr.mgh.harvard.edu>

of 81 spectral power features were computed, corresponding to the nine frequency bands, per seven electrode locations plus the two virtual inter-hemispheric channels, i.e.,: F3, Fz, F4, Cz, P3, Pz, P4, F1-F3, and P3-P4.

Coherence

Coherence features measure the co-variance related to amplitude and phase between a pair of electrodes in each frequency band. In this work, we used the magnitude squared coherence (MSC), which measures the co-variance between two power spectra. To calculate such features, we utilized the Hilbert transform to compute the analytic signals $x(k)$ and $y(k)$ for two electrodes of a pair in an epoch for a given frequency band. Then, the calculation of the auto-spectrum and cross-spectrum were computed by Equations 3.2 and 3.3, respectively:

$$P_{xx} = \frac{1}{m} \sum_{k=1}^m x(k)\tilde{x}(k), \quad (3.2)$$

$$P_{xy} = \frac{1}{m} \sum_{k=1}^m x(k)\tilde{y}(k). \quad (3.3)$$

Here, $\tilde{x}(k)$ and $\tilde{y}(k)$ are the conjugate of the analytic signals $x(k)$ and $y(k)$, respectively. Then, the MSC coherence can be defined as::

$$MSC = \frac{|P_{xy}|^2}{P_{xx}P_{yy}}. \quad (3.4)$$

For coherence measurements, the following frequency bands were considered: delta (0.1-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), beta (12-30 Hz), and gamma (30-100 Hz). The computation totalled 25 coherence features, derived from five frequency bands for five electrode connections, namely : Fz-Pz, F3-F4, P3-P4, F3-P3, and F4-P4.

Amplitude modulation rate-of-change

Figure 3.2 depicts the signal processing steps involved in the computation of the amplitude modulation rate-of-change features proposed in [38]. To compute the amplitude modulation fea-

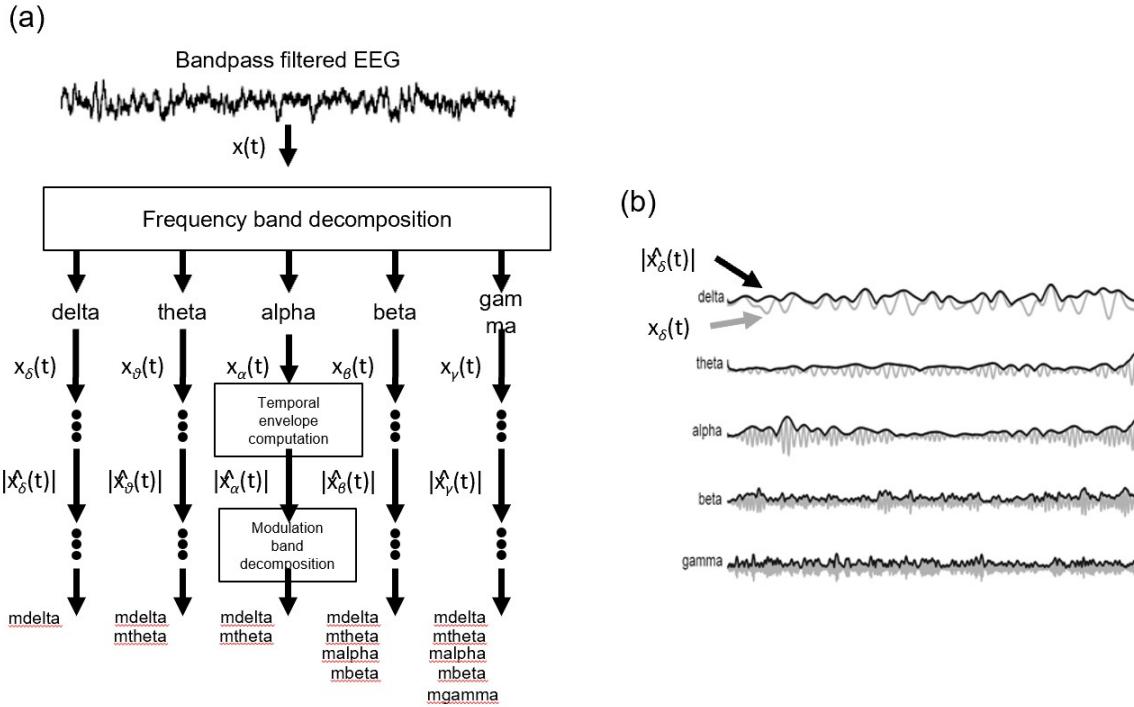


Figure 3.2 – (a) Signal processing steps involved in computation of amplitude modulation rate-of-change features. (b) Bandpass filtered EEG signals and their time envelopes. Adapted from [38]

tures, the EEG signal is first decomposed in the delta, theta, beta, and gamma bands. Next, the Hilbert transform is applied to compute each subband envelope. A second frequency decomposition is then applied to the band envelope signals to identify the changes in their amplitude modulation. Lastly, the average energy per frequency-modulation band is computed, normalized by the total signal energy, and used as a feature. Given properties of the Hilbert transform, not all frequency-modulation frequency combinations are possible, thus the following combinations are used here: delta-mdelta, theta-mdelta, theta-mtheta, alpha-mdelta, alpha-mtheta, beta-mdelta, beta-mtheta, beta-malpha, beta-mbeta, gamma-mdelta, gamma-mtheta, gamma-malpha, gamma-mbeta, and gamma-mgamma. Hence, a total of 126 features were extracted (14 frequency-modulation bands x 9 electrodes) to quantify amplitude modulation spectra rate-of-change.

Modulation frequency “patches”

Recently, the work in [19] showed that improved AD diagnostics could be achieved if non-conventional bands were used in the calculation of the above mentioned amplitude modulation

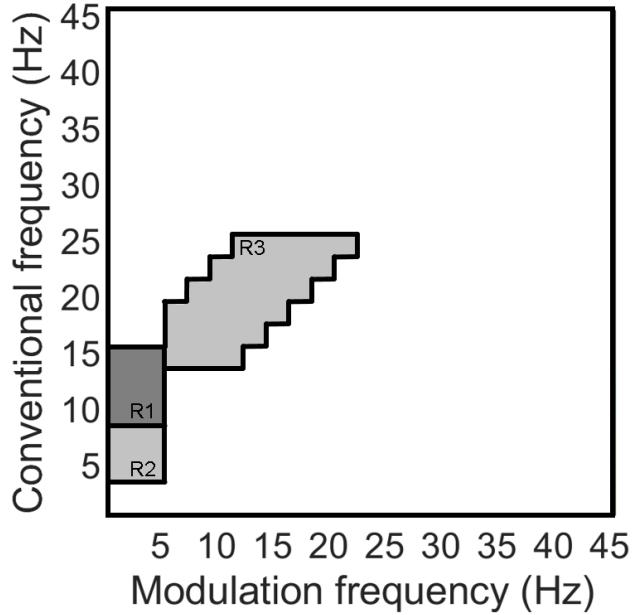


Figure 3.3 – Regions identified in the modulation power domain. Image taken from [20].

rate-of-change features. These so-called “patches,” as seen in Fig. 3.3, showed to be important in discriminating mild cognitive impairment from AD, as well as moderate AD from severe AD. A total of three patches (termed R1-R3 in the figure) and their ratios showed to provide the most discriminatory information. Here, a total of 54 features are extracted corresponding to the three patches and their ratio combinations for each of the seven electrodes and two virtual bipolar channels.

Overall, a total of 286 features were extracted from each of the 8-s epochs. Table 3.2 summarizes the features calculated for each EEG epoch.

EEG feature functionals

As mentioned above, EEG features are computed per time epoch. MRI features, in turn, are available per subject. As such, in order to enable feature-level fusion experiments, we must aggregate all EEG features into one final per-subject feature. Here, summary statistics are used as “functionals” for the purpose of temporal feature aggregation. More specifically, mean, standard deviation, coefficient of variation, median, skewness, and kurtosis were obtained for each of the four

Table 3.2 – Summary of the features computed from EEG signals.

Features Group	Features
Spectral power	delta theta alpha low-alpha high-alpha beta delta-beta theta-beta gamma
Coherence	MSC: delta MSC: theta MSC: alpha MSC: beta MSC: gamma
Amplitude modulation rate-of-change	delta-mdelta theta-mdelta theta-mtheta alpha-mdelta alpha-mtheta beta-mdelta beta-mtheta beta-malpha beta-mbeta gamma-mdelta gamma-mtheta gamma-malpha gamma-mbeta gamma-mgamma
Modulation frequency “patches”	R 1 R 2 R 3 R 1 /R 3 R 2 /R 1 R 2 /R 3

EEG feature types (spectral power, coherence, amplitude modulation, and modulation region) [32].

More details are provided below:

1. Mean: The mean value is determined by summing each of the samples and dividing it by the total number of observations. It is considered a measure of central tendency, since it tries to describe the distribution by a central value.

2. Standard deviation: The standard deviation (std) measures how the samples of the distribution spread around its mean value. It is defined as the square root of the variance, which is determined by the average squared distance between each quantity and the mean.
3. Coefficient of variation: The sample coefficient of variation (cv) is established as the ratio of the standard deviation to the mean of a distribution. It is useful supplementary information to the standard deviation in describing the variability of the distribution, since it allows the comparison of data with different mean values.
4. Median: Median is also a measure of central tendency. The median value is designated by the value of the sample standing in the middle of the distribution once it has been ordered from its smallest value to its highest value. It can also be interpreted as the value that segregates the lower half of the samples from the higher half.
5. Skewness: Skewness indicates a measure of the degree of symmetry of the probability distribution concerning its mean. The measure could be used to suggest how different a distribution is from a normal distribution. If the data is accumulated to the left, which forms a tail in the right direction, the skewness value will be positive. On the contrary, if the data is piled high into the right, forming a tail that points to the left, a negative value will be attributed to the skewness. The skewness can also be defined as zero, which suggests that it is a perfect normal distribution.
6. Kurtosis: Kurtosis is a measure that refers to the presence of outliers in the data by comparing how tailed the distribution is concerning a normal distribution. A high value of kurtosis indicates a heavy tail caused by the presence of multiple outliers. Moreover, a low value of kurtosis shows a tendency of light tail, indicating the shortage of outliers.

In the end, a total of 1716 EEG features exist corresponding to the 286 per-epoch features times their 6 summary statistics.

3.4.2 MRI feature extraction

MRI features were extracted with the utilization of Freesurfer software. Subcortical volumes, cortical thickness (CT), and area were extracted from different regions-of-interest (ROIs). Features

were extracted by the group of Dr. William J. McGeown from Strathclyde University as part of an ongoing collaboration with Cognision, and more detail can be found in as following described [75].

Subcortical segmentation features

Subcortical segmentation of the normalized brain volume was performed to account for head size variation using the automatic subcortical segmentation of a brain volume provided by the Freesurfer software. Segmentation based on the ASEG atlas was used, which locates the structure by probabilistic means [43]. Volumes of anatomical structures as lateral ventricle, inferior lateral ventricle, cerebellum white matter, cerebellum cortex, thalamus proper, caudate, putamen, pallidum, hippocampus, amygdala, accumbens area, ventral DC, vessel, choroid plexus, cortex, cerebral white matter, and surface holes were determined for the left and right hemispheres.

Moreover, the 3rd ventricle, fourth ventricle, fifth ventricle, cerebrospinal fluid (CSF), WM hypointensities, non-WM hypointensities, optic chiasm, posterior corpus callosum, midposterior corpus callosum, central corpus callosum, midanterior corpus callosum, anterior corpus callosum, brain segmentation, brain segmentation without ventricles, brain segmentation without ventricles from surf, cortex, cerebral white matter, subcortical gray matter, total gray matter, supratentorial, supratentorial without ventricles, supratentorial without ventricles voxel, mask, ratio of brain segmentation to estimated total intracranial volume (eTIV), ratio of mask to eTIV, total number of defect holes in surfaces before fixing, and estimated total intracranial volumes features were also measured. A total of 62 features were gathered for subcortical segmentation measurements.

Cortical parcellation

Average thickness and area of the cortical parcellation were also calculated with Freesurfer. MRI scans were parcellated into distinct anatomical regions, according to the Desikan-Killiany atlas [44]. Anatomical cortical structure measurements were determined for the left and right hemispheres for bankssts, caudal anterior cingulate, caudal middle frontal, cuneus, entorhinal, fusiform, inferior parietal, inferior temporal, isthmus cingulate, lateral occipital, lateral orbitofrontal, lingual, medial orbitofrontal, middle temporal, parahippocampal, paracentral, pars opercularis, pars orbitalis, pars triangularis, pericalcarine, postcentral, posterior cingulate, precentral, precuneus, rostral ante-

Table 3.3 – Summary of the MRI computed features.

Features Category	Features
Subcortical segmentation - ASEG	62 volumes measurements of anatomical provided by the subcortical segmentation.
Cortical parcellation - APARC	148 thickness and area measurements of anatomical structures resulted from the cortical parcellation.
White matter parcellation - WM	75 thickness measurements of anatomical structures resulted from the white matter parcellation.

rior cingulate, rostral middle frontal, superior frontal, superior parietal, superior temporal, supramarginal, frontal pole, temporal pole, transverse temporal, insula, mean thickness, thickness from brain segmentation volume without ventricles, thickness from eTIV, white surface area, area from brain segmentation volume without ventricles, and area from eTIV. A total of 148 features were computed for cortical parcellation measurements.

White matter parcellation features

White matter parcellation also followed the regions according to the Desikan- Killiany atlas. Therefore, anatomical cortical structure measurements were estimated for the left and right hemispheres for the following ROI's: bankssts, caudal anterior cingulate, caudalmiddle frontal, cuneus, entorhinal, fusiform, inferior parietal, inferior temporal, isthmus cingulate, lateral occipital, lateralarbitofrontal, lingual, medial orbitofrontal, middle temporal, parahippocampal, paracentral, pars opercularis, pars orbitalis, pars triangularis, pericalcarine, postcentral, posterior cingulate, precentral, precuneus, rostral anterior cingulate, rostral middlefrontal, superior frontal, superior parietal, superior temporal, supramarginal, frontal pole, temporal pole, transverse temporal, insula, unsegmented white matter, and cerebral white matter. The entire cerebral white matter volume, mask volume, and estimated total intracranial volume were determined. There were a total of 75 features resulting from the white matter parcellation process.

Table 3.3 summarizes the 285 features calculated for each of the MRI scans. Subsequently, all the calculated features were submitted to a normalization process. Normalization was utilized to scale all the features in a comparable range from the value of -1 to 1 and avoid future problems with regression models that could be dependent on distance calculations.

3.5 Feature selection

At the end of the feature extraction procedure, a total of 1716 features were calculated from the EEG signals and a total of 285 features were calculated from MRI scans. While our main goal is to explore the benefits of a multimodal system, one exploratory aim we have is to also gauge the benefits of one single modality over the other, in particular the benefits of the newly proposed modulation spectrum patch features [20]. In the multimodal system, we use the feature fusion technique to combine modalities. In our experiments, feature fusion is achieved by concatenating the feature matrices into one final larger feature matrix.

Given the large number of features and the limited size of the available dataset, feature selection is needed to reduce the number of features to avoid overfitting [59]. Feature selection or ranking aims to rank available features based on their potential impact on the downstream classification task [15]. This way, irrelevant and/or redundant features are removed.

Many feature selection techniques have been reported in the literature and there is no consensus on what technique is ideal for a specific application, hence different methods are commonly explored. Here, different univariate filter techniques for feature selection are explored. Filter techniques do not utilize any learning model during the feature selection process, thus the ranking is based solely on data characteristics [6]. Combined with an univariate approach, the feature selection step evaluates each feature in the dataset one-by-one by determining their relationship with the dependent variable, or label to be predicted; in our case, MMSE scores. In the end, the features are ranked based on this relationship and the features with the highest scores are kept selected. In this study, three different filter features techniques were explored, as detailed next.

3.5.1 Correlation-based ranking

Correlation-based feature selection relies on correlation measures between the tested features and the outcome variables. Different types of correlations can be used to measure different associations between the two variables. Here, two of the most popular indices of correlation are used, namely Pearson and Spearman correlations. Pearson correlation ranges from -1 to 1 and measures a linear relationship between the variables, with a correlation coefficient of zero indicating no linear relationship. Spearman correlation, in turn, measures the ranking relationship between variables,

thus considers linear relationships and non-linear monotonic relationships between variables [27]. As mentioned in [18], correlation based methods are widely used in EEG-based AD diagnostic systems.

3.5.2 Minimum redundancy maximum relevance, MRMR

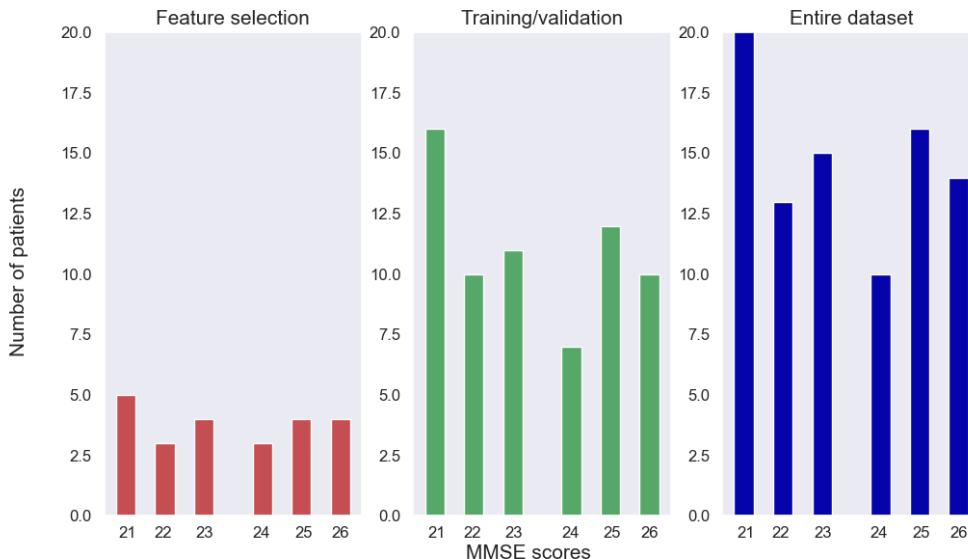
One main drawback of correlation-based methods is that while keeping relevant features, their redundancy is not accounted for, thus multiple features may be kept while providing limited additional information for the downstream classification task. As such, it may be seen as suboptimal. To compensate for this limitation, the minimum redundancy maximum relevance (MRMR) feature selection algorithm has been proposed. In essence, the algorithm selects a subset of features having the most relation with a class (relevance) and the least relation between themselves (redundancy). Relevance can be calculated with an F-statistic metric or mutual information, whereas redundancy can be calculated by using Pearson correlation coefficient or mutual information [15], [80], [69]. For our usage, the open-source implementation of the MRMR provided by [69] was used.

3.5.3 Top-features grouping

In an attempt to match the dimensionality reported in previous EEG-based AD literature (e.g., [38, 47, 23]), the top 24 features from each of the feature modalities was selected. Here, we call feature “Group 1” the top-24 features selected for the EEG modality. Group 2, in turn, corresponds to the top-24 MRI features. As we are particularly interested in multimodal systems, two other feature fusion strategies are explored. First, we apply feature selection to the fused EEG-MRI features sets and keep the top 24 features of this combined set. This is termed feature Group 3. Earlier work has suggested some correlation between EEG and MRI features [75], thus performing feature selection of the fused set may help remove redundant features from varying modalities. Lastly, we combine the top-12 EEG and top-12 MRI features and call this feature Group 4. Table 3.4 summarizes the four top feature Groups explored herein.

Table 3.4 – Top-feature grouping classes.

Groups	Description
Group 1	Top 24 features from the ranking of features from EEG signals.
Group 2	Top 24 features from the ranking of features from MRI scans.
Group 3	Top 24 features from the ranking of the combination of features from EEG and MRI.
Group 4	Combination of Top 12 features from group 1 with the top 12 features from group 2.

**Figure 3.4 – Distributions of the feature selection-train-test data partitions used.**

3.6 Data partitioning

From the data available from the 89 patients, 25% were exclusively dedicated to the feature selection procedure, while the remaining 75% of the data was used for training and testing of the regression models. Using disjoint sets for feature selection and classifier training reduces any unwanted biases in the reported performance figures. Data partitioning was done randomly while keeping the distribution of the MMSE scores proportional to the entire dataset. The partitioning can be summarized in Figure 3.4. This partitioning was done five times in order to assure that the selected features are not overly sensitive to data partitioning.

3.7 Classification/Regression

Since we are interested in AD severity monitoring, a regression task is needed. More specifically, we are interested in estimating the MMSE score of an AD patient. Motivated by insights from [18], three supervised regressors are explored [51]: Support vector machine regression with linear (SVM-Linear) and Gaussian (SVM-RBF) kernels, random forest regression (RF), and k-nearest neighbours regression (KNN). All the models were implemented with the open-source scikit-learn library for python [79]. More details are provided next.

3.7.1 Support vector machine

The support vector machine regression (SVMR) model searches for a regression function in which all the obtained target errors will be under a specific value [94]. The strategy used in the implementation of the SVMR is very similar to the strategies used by the support vector machine classifier. In the regression scenario, the optimal hyperplane will not be used to separate the two sets of data but to map the training data into a regression function, and the margin in this scenario is used as a tolerance margin in which the errors must be under [84].

3.7.2 Random forest

Random Forest (RF), in turn, combines many decision trees into a single model. The structure of a random forest is composed of several decision trees running in parallel with no interaction between them. Ultimately, the random forest outputs the mean value of each tree outputs that composes the random forest [68].

3.7.3 K-nearest neighbour

The k-nearest neighbour classifier (KNN) assumes that an unseen sample likely belongs to the same class as the k most analogous-distant neighbours [82]. For regression, the algorithm takes the weighted average of the k nearest neighbours, weighted by the inverse of their distance.

3.8 Evaluation metric

To evaluate the performance of each model, we use the root mean squared error (RMSE) between the prediction and the true MMSE score as a metric. Since the dataset available is not very large, we rely on repeated k-fold cross-validation as a testing setup. In this way, the available train/test dataset (i.e., 75% of the original dataset) is randomly split into k folds (per repetition). Training is done with k-1 folds and then tested on the held-out fold. This is repeated until all folds have been used as hold-out. With repeated cross-validation, this procedure is repeated N times, each time with a different partition. In the end, a total of $N \times K$ RMSE values are produced. Here, $N = 50$ and $K = 5$ are used. The average and standard deviation across these $N \times K$ values are used to gauge system performance. Ideally, we are interested in models that achieve a low average RMSE, as well as a low standard deviation, thus suggesting the model is insensitive to data partitioning.

Chapter 4

Results and discussion

In this chapter, the efficiency of the proposed models and the extracted features is explored. The advantages of a multimodal system over systems relying on a single modality are also highlighted and discussed.

4.1 Statistical analyses

As a first step, we wanted to assess the relationship between MMSE scores and the EEG and MRI features extracted herein. To this end, Spearman rank correlation was used and only features that had a p-value smaller than 0.05 were retained for analysis. Tables 4.1 and 4.2 show the Spearman correlation coefficients for the EEG and MRI features, respectively. As can be seen, 39 EEG features met the significance criterion, whereas 29 MRI features were found. Absolute correlation values ranged from 0.21-0.31 for EEG features and from 0.21-0.37 for MRI, thus suggesting that individually, MRI may provide more relevant details for AD progression. Clearly, EEG amplitude modulation features showed to achieve the greatest correlation, while cortical parcellation and white matter parcellation features showed to be the best for MRI. From the ten top features from EEG, three corresponded to the new modulation spectral patch features, thus suggesting that may indeed be useful for the task at hand.

Since prior work has also suggested a correlation between EEG and MRI features [75], here we explore the relationship between feature modalities. Figure 4.1 depicts a heatmap according to the

Table 4.1 – Ranking of the EEG features correlated with MMSE scores. All the values presented a p-value smaller than 0.05.

Features	Correlation	p-value
kurtosis-gamma-mgamma-F4	0.314753	0.002663
skewness-R2oR3-Pz	0.289395	0.005947
std-alpha-mdelta-P3-P4	-0.27353	0.009497
kurtosis-beta-mbeta-F4	0.258975	0.014262
skewness-msc-delta-Fz-Pz	0.250721	0.017796
std-R1-P3-P4	-0.24512	0.020604
skewness-R2oR1-F3-F4	-0.2449	0.020722
median-tab-P3-P4	0.243892	0.021265
median-delta-P3-P4	-0.24096	0.022923
kurtosis-gamma-mtheta-Cz	0.233574	0.027598
mean-delta-mdelta-P3-P4	-0.23126	0.029217
mean-delta-P3-P4	-0.22935	0.030619
cv-delta-mdelta-P3-P4	-0.22878	0.031046
median-R1-P3-P4	0.227577	0.031966
mean-R1-P3-P4	0.227171	0.032282
skewness-R2oR3-Cz	0.226825	0.032553
skewness-R2oR3-P4	0.226323	0.032949
mean-beta-mdelta-P4	0.22615	0.033087
skewness-alpha2-P4	-0.22573	0.033427
std-dtab-P3-P4	0.224863	0.034126
mean-tab-P3-P4	0.224135	0.034726
median-delta-mdelta-P3-P4	-0.22279	0.035858
median-beta-mdelta-P4	0.22153	0.036946
mean-beta-mdelta-Pz	0.220578	0.037786
skewness-theta-P4	0.220198	0.038125
std-R3-P3-P4	-0.21933	0.038909
cv-beta-mdelta-F4	0.21768	0.040443
mean-beta-mdelta-P3	0.217438	0.040671
median-beta-mdelta-P3	0.217143	0.040951
median-beta-mdelta-P3-P4	0.215215	0.042823
mean-beta-mdelta-P3-P4	0.214773	0.043262
std-msc-theta-F3-P3	0.214396	0.043639
kurtosis-gamma-mgamma-Pz	-0.21341	0.044643
skewness-R2oR3-Fz	0.213319	0.044732
kurtosis-alpha1-F3-F4	0.212182	0.04591
median-msc-gamma-F4-P4	-0.21168	0.046435
cv-delta-P3-P4	-0.21133	0.046811
skewness-alpha2-Pz	-0.21103	0.047134
std-tab-P3-P4	-0.20875	0.049629

Table 4.2 – Ranking of the MRI features correlated with MMSE scores. All the values presented a p-value smaller than 0.05.

Features	Correlation	p-value
aparc_CT_rh_fusiform_thickness	0.372915	0.000319
aparc_area_lh_lateraloccipital_area	0.329514	0.001615
aparc_CT_rh_insula_thickness	0.328219	0.001689
aparc_area_rh_lateraloccipital_area	0.309018	0.003214
wm_rh_lateraloccipital	0.308307	0.003288
wm_rh_fusiform	0.292803	0.00536
wm_lh_lateraloccipital	0.257521	0.014836
aparc_CT_rh_parahippocampal_thickness	0.249052	0.018595
aparc_area_rh_inferior temporal_area	0.247406	0.019414
wm_lh_lingual	0.247399	0.019417
aparc_area_rh_bankssts_area	0.246363	0.019948
wm_rh_precuneus	0.24636	0.019949
aseg_Right_choroid_plexus	-0.24501	0.020659
aparc_area_lh_inferior temporal_area	0.240132	0.023412
aparc_CT_rh_temporalpole_thickness	0.234593	0.026909
aparc_CT_lh_parahippocampal_thickness	0.233237	0.02783
aparc_area_rh_middle temporal_area	0.230571	0.029718
aparc_area_rh_transverse temporal_area	0.230544	0.029737
wm_rh_parsorbitalis	0.22978	0.030298
aparc_CT_rh_rostralmiddlefrontal_thickness	-0.22607	0.03315
wm_rh_transverse temporal	0.22602	0.033191
aparc_area_rh_precuneus_area	0.218985	0.039228
wm_rh_cuneus	0.217247	0.040852
aparc_area_lh_rostralmiddlefrontal_area	0.213918	0.044122
wm_rh_lingual	0.213458	0.04459
wm_rh_medialorbitofrontal	0.212549	0.045527
aparc_area_lh_frontalpole_area	0.209892	0.048362
aseg_Right_Amygdala	0.208465	0.049943
aparc_CT_rh_bankssts_thickness	0.208446	0.049965

Spearman correlation coefficient across all features of EEG and MRI modalities. The EEG features are represented by the bigger black rectangle below the EEG label, and it shows the correlation coefficient among the EEG variables. In turn, the smaller black rectangle above the MRI label shows the correlation associated with the features extracted from the MRI scans. The areas that do not belong to any of the black rectangles represent the correlation among the EEG and MRI features. As can be seen, some relationship between modalities exists, but inter-feature correlations within the same modality are substantially higher, suggesting that feature Groups 1, 2, and 4 may lead to better results.

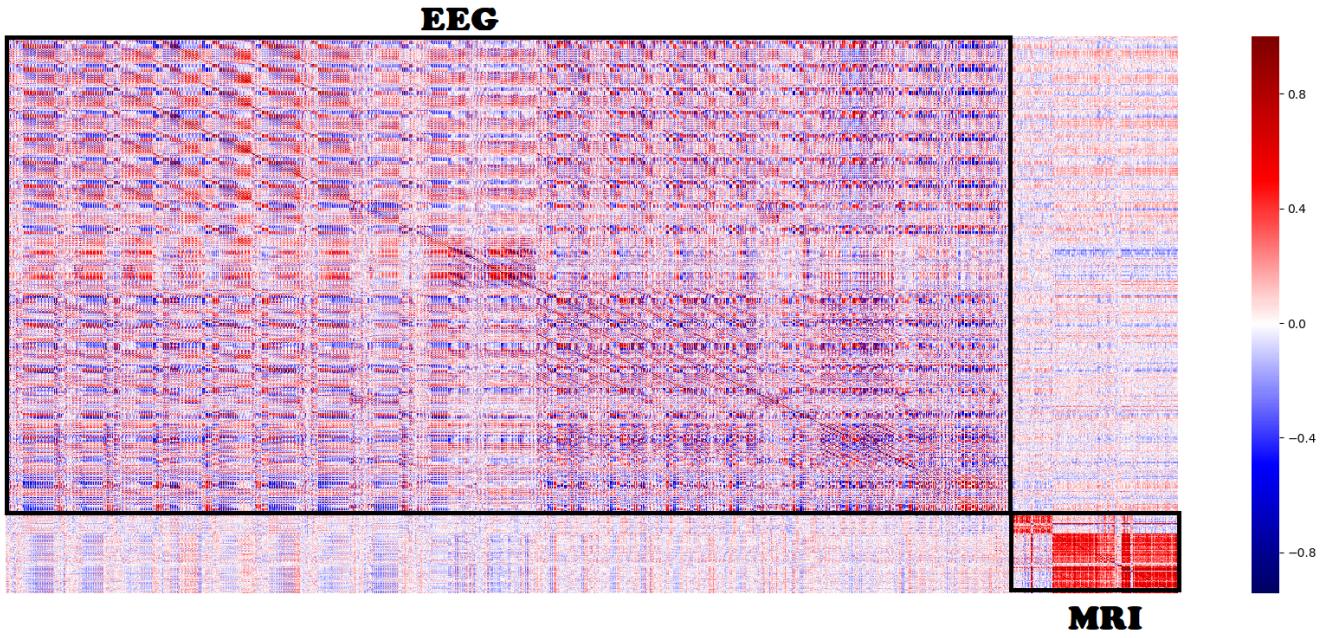


Figure 4.1 – Relationship among the EEG and MRI extracted features.

4.2 AD severity prediction

Here, we explore the benefits of AD severity prediction using unimodal and multimodal systems. Experimental results for the four feature groups are shown in Table 4.3, where the average RMSE score followed by its standard deviation is reported over all the cross-validation trials and averaged over the five test setup runs. Moreover, Figures 4.2, 4.4, and 4.3 present the distributions of the overall average RMSE scores based on the different regression algorithms, different groups of features, and the different feature selection algorithms, respectively.

Additionally, with the intention of investigate whether the observed differences in performance are statistically significant, a one-way ANOVA statistical test was performed for different populations of RMSE for each of the divisions as seen in Figures 4.2, 4.4, and 4.3. All the results were significant at $p < 0.05$.

As can be seen from Table 4.3 and Figure 4.2, the results achieved suggest that SVM-RBF and random forest classifier achieve lower prediction errors relative to an SVM with a linear kernel or a k-NN regressor. The SVM-RBF classifier, in turn, achieved a lower RMSE variability, suggesting the

Table 4.3 – Performance comparison across regressors, feature selection algorithms, and feature groups.

MRMR				
	Group 1	Group 2	Group 3	Group 4
SVM Linear	2.065 ± 0.269	1.847 ± 0.208	2.018 ± 0.245	1.931 ± 0.244
SVM RBF	1.849 ± 0.198	1.833 ± 0.196	1.845 ± 0.198	1.830 ± 0.194
Random Forest	1.897 ± 0.231	1.928 ± 0.226	1.951 ± 0.238	1.881 ± 0.228
KNN	2.009 ± 0.261	1.894 ± 0.242	2.017 ± 0.276	2.011 ± 0.265
Pearson				
	Group 1	Group 2	Group 3	Group 4
SVM Linear	2.021 ± 0.261	1.859 ± 0.218	1.958 ± 0.259	1.924 ± 0.246
SVM RBF	1.843 ± 0.195	1.828 ± 0.195	1.835 ± 0.194	1.834 ± 0.194
Random Forest	1.963 ± 0.231	1.844 ± 0.247	1.889 ± 0.257	1.840 ± 0.245
KNN	2.057 ± 0.256	1.971 ± 0.258	2.010 ± 0.272	1.949 ± 0.248
Spearman				
	Group 1	Group 2	Group 3	Group 4
SVM Linear	2.019 ± 0.255	1.864 ± 0.209	1.968 ± 0.257	1.920 ± 0.245
SVM RBF	1.848 ± 0.194	1.830 ± 0.195	1.841 ± 0.193	1.834 ± 0.193
Random Forest	1.917 ± 0.223	1.849 ± 0.245	1.854 ± 0.238	1.827 ± 0.240
KNN	2.031 ± 0.253	1.952 ± 0.251	2.009 ± 0.265	1.954 ± 0.257

model is less sensitive to data partitioning, an important factor for system generality. Moreover, the results from the Table and Figure 4.3 suggest that feature selection algorithms do not play a crucial role (more on top features in the following section) on final performance. Regarding feature groups (see Figure 4.4), between unimodal systems (Group 1 and Group 2), MRI features achieved somewhat lower RMSE values relative to EEG features, thus corroborating recent findings [39]. Overall, for most cases tested, multimodal systems, particularly trained on feature Group 4, showed slight improvements in performance relative to unimodal systems. In particular, lower RMSE variability was seen, thus suggesting that multimodal systems are less sensitive to data partitioning. While a random forest regressor with Group 4 features resulted in the lower average RMSE (1.827), the variability was rather large. For practical applications, however, a combination of lower RMSE average and variability is desirable. As such, the SVM-RBF method, combined with either Group 3 or Group 4 features, showed to be the most promising solutions from purely a performance point of view.

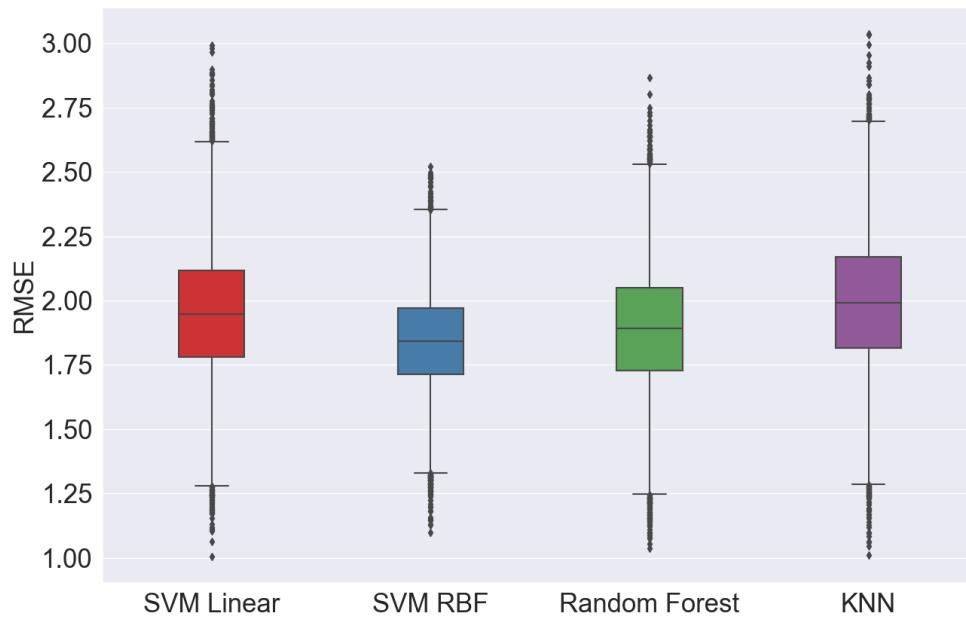


Figure 4.2 – Distribution of the RMSE scores according to regressors used.

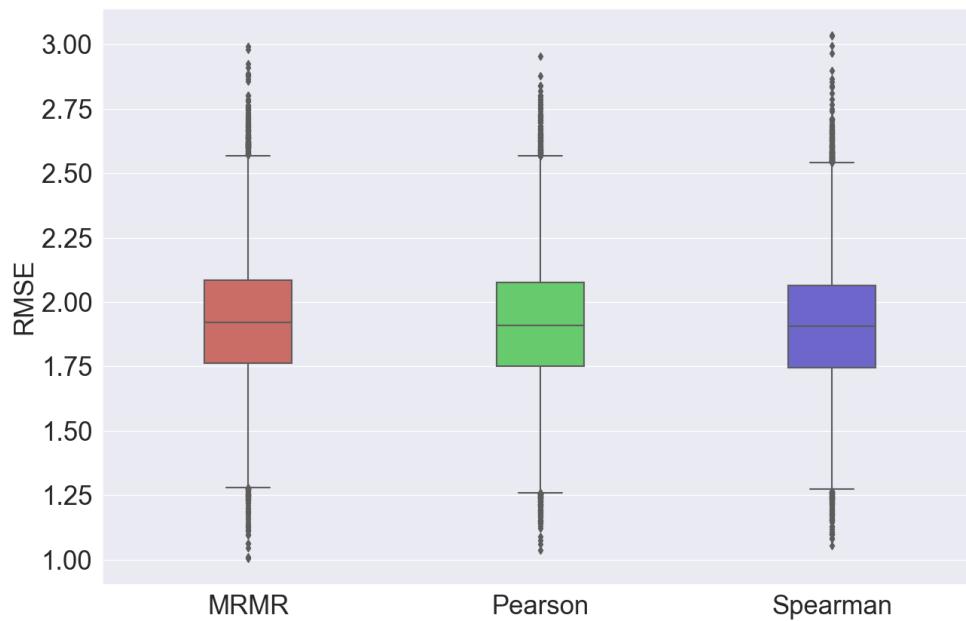


Figure 4.3 – Distribution of the RMSE scores according to the features selection techniques.

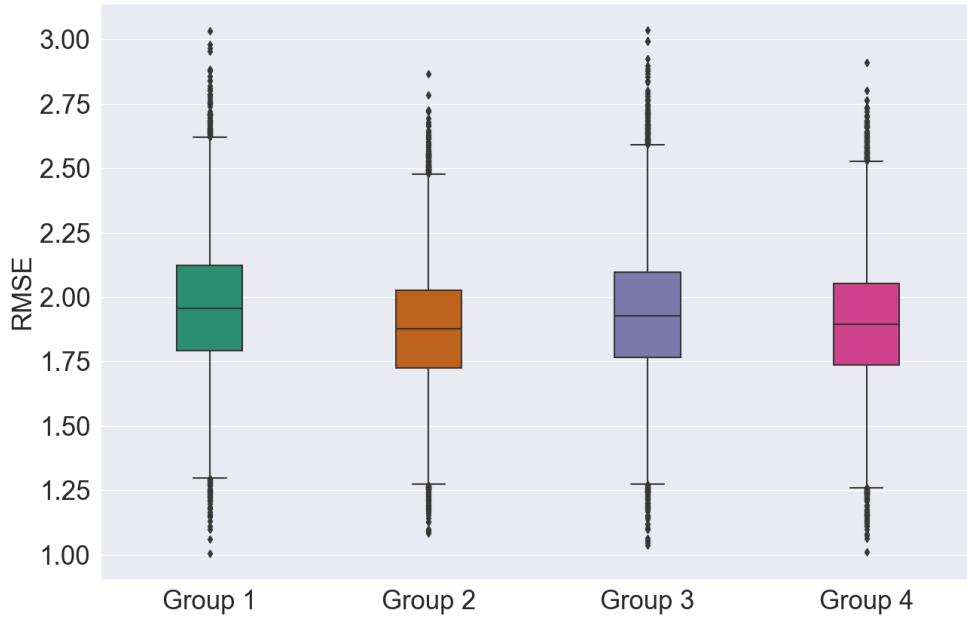


Figure 4.4 – Distribution of the RMSE scores according to the groups of features.

4.3 Consistently selected features

In order to better understand the impact of different feature modalities on overall severity monitoring performance, an in-depth analysis is done concerning the features that were consistently selected during the five test setup runs. To this end, Tables 4.4, 4.5, and 4.6 show the list of the features which were selected more than once by each of the feature selection techniques for feature Groups 1-3, respectively.

From the Tables, it can be seen that MRMR demonstrated to be less sensitive to data partitioning than the other feature selection techniques, since it was able to select similar features across the five runs consistently. This was true for all feature Groups tested. This is an important aspect that needs to also be taken into account when evaluating overall system performance [62]. This, combined with the improved performance achieved with the SVM-RBF regressors, suggests that the most reliable system for AD severity detection would be a system based on MRMR feature selection, and SVM-RBF regressor and on a multimodal feature set (Group 4).

Table 4.4 – Consistently Selected Features from Group 1 - EEG Features.

EEG	MRMR	Pearson	Spearman
kurtosis-R2oR3-P3-P4	5	0	0
skewness-beta-mdelta-F4	4	0	0
cv-theta-mdelta-P3	4	0	0
cv-R2oR3-P3-P4	3	0	0
skewness-delta-mdelta-P4	2	0	0
skewness-gamma-mgamma-F3-F4	2	0	0
skewness-theta-mdelta-P3	2	0	0
skewness-gamma-mtheta-P3-P4	2	0	0
skewness-R2oR3-F3	2	0	0
skewness-R3-F3-F4	2	0	0
skewness-alpha2-Pz	2	0	0
kurtosis-gamma-mgamma-F3-F4	2	0	0
skewness-alpha-F3	2	0	0
skewness-gamma-mbeta-F4	2	0	0
skewness-R2oR1-P3-P4	2	0	0
skewness-R2oR3-P3-P4	2	0	0
median-R2oR3-P3-P4	2	0	0
std-R2oR3-P3-P4	2	0	0
mean-R2oR3-P3-P4	2	0	0
kurtosis-R2oR1-P3-P4	2	0	0
skewness-tab-F4	2	0	0
skewness-R1-F3	2	0	0
skewness-msc-beta-Fz-Pz	2	0	0
skewness-R2oR3-F3-F4	2	0	0
kurtosis-beta-mtheta-Cz	0	2	2
median-msc-beta-P3-P4	0	2	0
mean-msc-beta-P3-P4	0	2	0
skewness-beta-mbeta-P4	0	2	0
skewness-R2oR3-Pz	0	2	2
skewness-R2oR1-P4	0	2	2
kurtosis-alpha1-F3-F4	0	2	2
skewness-R2oR3-Cz	0	2	0
skewness-msc-delta-F3-P3	0	0	2
std-alpha-mdelta-P3-P4	0	0	2
cv-msc-beta-F3-P3	0	0	2
kurtosis-R1-F3-F4	0	0	2

From the analyses of Table 4.4, features *kurtosis – R2oR3 – P3 – P4*, *skewness – beta – mdelta – F4*, *cv – theta – mdelta – P3*, and *cv – R2oR3 – P3 – P4* showed to be the features most often selected, thus highlighting the importance of modulation spectral features not only for diagnostics, but also for AD severity monitoring. The virtual electrode *P3 – P4* also showed to be important, thus corroborating the inter-hemispheric deficit commonly reported with AD [66]. More importantly, from the newly-proposed modulation spectral patch features available, they represented half of the top-20 EEG features selected by MRMR, thus highlighting their importance for unimodal EEG-based monitoring systems.

From Table 4.5, in turn, a larger number of features derived from MRI scans were consistently selected, relative to Group 1. Areas measured from the left frontal pole and superior temporal were selected in all the runs by the MRMR technique, thus corroborating the importance of frontal and temporal lobes on AD diagnosis [61]. Moreover, while AD is mostly associated with its effects on the cortical grey matter, damages to the white matter have also been reported with AD. The authors of [13], for example, found differences in the scans of patients with AD from healthy controls in the white matter of the frontal, temporal, and parietal lobes. Here, MRMR selected white matter mask volume, cerebral white matter volume, and estimated total intracranial volume in four of the five runs. Right side measurements from the frontal pole, supramarginal, superior temporal, superior parietal, transverse temporal, cerebral white matter vol, unsegmented white matter, and insula were also selected in four out of the five runs, aside with the left side measurements of the cerebral white matter vol, and the entorhinal area. Additionally, areas from the left and right side of lateral occipital, and white matter right-side measurements of the lateral occipital and fusiform were selected at least four times from Pearson and Spearman correlation feature selection technique. These are all areas and features that have been previously linked to AD (e.g., [39, 67, 12]).

Lastly, Table 4.6 displays the consistently-selected features from Group 3. Of the 29 features that were selected in more than one run, only 6 were from MRI. The top features also showed up previously during the unimodal tests, thus suggesting their importance for the task at hand.

Tables 4.7 and 4.8 show the overall summary of consistently selected features based on feature category, frequency bands, electrode positions, and brain hemispheres. From the EEG features, the skewness features stood out in the unimodal systems, whereas the standard deviation features stood out for multimodal ones. Such findings suggest that feature temporal dynamics play an im-

Table 4.5 – Consistently Selected Features from Group 2 - MRI Features.

MRI	MRMR	Pearson	Spearman
aparc_area_lh_frontalpole_area	5	0	0
aparc_area_lh_superiortemporal_area	5	0	0
wm_MaskVol	4	0	0
wm_CerebralWhiteMatterVol	4	0	0
wm_rh_frontalpole	4	0	0
wm_rh_supramarginal	4	0	0
wm_rh_superiortemporal	4	0	0
wm_rh_superiorparietal	4	0	0
wm_rh_transversetemporal	4	0	0
aparc_area_lh_entorhinal_area	4	2	2
wm_rhCerebralWhiteMatterVol	4	0	0
wm_Right_UnsegmentedWhiteMatter	4	0	0
wm_rh_insula	4	0	0
wm_EstimatedTotalIntraCranialVol	4	0	0
wm_lhCerebralWhiteMatterVol	4	0	0
aparc_area_lh_lateraloccipital_area	0	4	3
wm_rh_lateraloccipital	0	3	4
wm_rh_fusiform	0	3	4
aparc_area_rh_lateraloccipital_area	2	4	3
aparc_CT_rh_fusiform_thickness	0	3	3
aparc_CT_rh_rostralmiddlefrontal_thickness	0	3	2
aparc_area_lh_postcentral_area	0	3	0
aseg_Left_Putamen	0	2	3
aseg_Right_Putamen	0	0	3
wm_rh_rostralmiddlefrontal	3	0	0
aseg_lhSurfaceHoles	3	0	0
wm_rh_postcentral	3	0	0
aparc_area_rh_transversetemporal_area	2	3	2
wm_rh_rostralanteriorcingulate	2	0	0
wm_lh_pericalcarine	2	0	0
wm_Left_UnsegmentedWhiteMatter	2	0	0
aseg_3rd_Ventricle	2	0	0
wm_rh_inferiortemporal	2	0	2
wm_rh_superiorfrontal	2	0	0
wm_rh_posteriorcingulate	2	0	0
wm_rh_bankssts	0	2	2
aparc_CT_rh_insula_thickness	0	2	2
aseg_SubCortGrayVol	0	2	2
wm_lh_lingual	0	2	2
aparc_area_lh_lingual_area	0	2	2
aparc_area_lh_fusiform_area	0	2	2
wm_lh_bankssts	0	2	2
wm_lh_entorhinal	0	2	2
aparc_CT_rh_temporalpole_thickness	0	2	2
aseg_Right_Amygdala	0	2	0
aparc_area_rh_inferiortemporal_area	0	2	0
aparc_area_rh_precuneus_area	0	2	0
aseg_Right_Hippocampus	0	2	0
wm_lh_inferiortemporal	0	2	0
aparc_CT_lh_parahippocampal_thickness	0	2	0
aparc_area_rh_fusiform_area	0	2	0
wm_lh_postcentral	0	2	0
aparc_area_rh_caudalmiddlefrontal_area	0	0	2
aparc_area_lh_rostralmiddlefrontal_area	0	0	2
aparc_area_lh_bankssts_area	0	0	2
wm_rh_parstriangularis	0	0	2

Table 4.6 – Consistently Selected Features from Group 3 - EEG and MRI Features.

EEG + MRI	MRMR	Pearson	Spearman
wm_EstimatedTotalIntraCranialVol	5	0	0
cv-alpha-mdelta-Cz	5	0	0
mean-alpha-mdelta-Cz	5	0	0
mean-gamma-mgamma-Fz	4	0	0
kurtosis-beta-mtheta-F3-F4	4	0	0
cv-beta-mtheta-Pz	3	0	0
std-gamma-mtheta-Pz	2	0	0
kurtosis-gamma-mdelta-Cz	2	0	0
std-alpha-P3-P4	2	0	0
mean-gamma-mbeta-Cz	2	0	0
std-msc-gamma-Fz-Pz	2	0	0
cv-alpha2-Fz	2	0	0
std-R1-P4	2	0	0
kurtosis-alpha1-F3-F4	0	2	2
kurtosis-beta-mtheta-Cz	0	2	2
wm_rh_lateraloccipital	0	2	2
skewness-R2oR1-P4	0	2	2
skewness-R2oR3-Cz	0	2	0
median-msc-beta-P3-P4	0	2	0
aparc_area_rh_lateraloccipital_area	0	2	0
mean-msc-beta-P3-P4	0	2	0
aparc_area_lh_entorhinal_area	0	2	0
aparc_CT_rh_fusiform_thickness	0	2	0
skewness-R2oR3-Pz	0	0	2
kurtosis-R1-F3-F4	0	0	2
cv-msc-beta-F3-P3	0	0	2
aseg_Left_Putamen	0	0	2
skewness-msc-delta-F3-P3	0	0	2
std-alpha-mdelta-P3-P4	0	0	2

portant role in severity monitoring; something that is corroborated by the importance seen with the modulation rate-of-change features. In fact, for both unimodal and multimodal systems, the amplitude modulation features showed to be extremely important, in particular features related to the beta frequency subband dynamics, thus corroborating previous research [56]. The *R2/R3* modulation patch regions, in turn, showed to be extremely important for unimodal systems. Regarding electrode positions, the virtual bipolar channels showed to be important for unimodal systems, thus in line with previous research showing an inter-hemispheric disconnection with AD [57]. From the MRI features, in turn, features derived from the APARC and WM groups were more often selected more than once than features from ASEG group. Moreover, features from the right side of the brain

Table 4.7 – Distribution of the selected features that were selected in multiple runs for EEG groups.

EEG		
Summary statistics		
	Unimodal	Multimodal
mean	2	4
standard deviation	2	5
coefficient of variation	3	4
median	2	1
skewness	21	4
kurtosis	6	5
Feature category		
	Unimodal	Multimodal
Spectral Power	4	3
Coherence	5	5
Amplitude Modulation	11	10
Modulation Domain	16	5
Feature subband / region		
	Unimodal	Multimodal
delta	2	1
theta	3	-
alpha	4	6
beta	7	6
gamma	4	5
R1	2	2
R3	1	-
R2/R3	10	2
R2/R1	3	1
Electrode position		
	Unimodal	Multimodal
P3-P4	12	4
F4	3	-
P3	2	-
P4	3	2
F3-F4	6	3
F3	3	-
Pz	3	4
Cz	2	6
F3-P3	2	2
Fz	-	2

showed up more often in the unimodal systems, thus corroborating right-dominance and laterality asymmetry findings from the literature (e.g., [70, 40]).

Table 4.8 – Distribution of the selected features that were selected in multiple runs for MRI groups.

MRI		
Feature category		
	Unimodal	Multimodal
APARC	20	3
ASEG	7	1
WM	29	2
Brain hemisphere		
	Unimodal	Multimodal
left	20	2
right	31	3

Table 4.9 – Comparative regression performance from models formed by the top features from Tables 4.4, 4.5 and 4.6

Model	EEG	MRI	EEG + MRI
SVM Linear	1.986 ± 0.224	1.745 ± 0.183	1.825 ± 0.203
SVM RBF	1.816 ± 0.157	1.807 ± 0.154	1.769 ± 0.149
Random Forest	1.798 ± 0.176	1.715 ± 0.186	1.682 ± 0.177
KNN	1.993 ± 0.217	1.887 ± 0.181	1.892 ± 0.209

As a last experiment, we build classifiers using only the consistently-selected features from the MRMR algorithm, as shown in Tables 4.4, 4.5 and 4.6 for EEG-only, MRI-only and EEG-MRI, respectively. Table 4.9 presents the results obtained from each of the four regressors in a cross-validation testing scheme. As previously, the SVM-RBF and random forest classifier achieved the best performance, thus further hinting at the importance of non-linear models. The SVM-RBF achieved the lowest variability across the cross-validation trials. Moreover, for unimodal systems, again MRI-based ones outperformed EEG-based ones. Notwithstanding, multimodal systems showed to always outperform at least one unimodal system. The best overall system was achieved with a multimodal system and a random forest regressor, achieving an average RMSE of 1.682.

Chapter 5

Conclusion

Alzheimer's disease affects millions of people worldwide, and the perspective is that this number will increase even more in the near future. Unfortunately, the disease is still incurable and irreversible, thus methods to diagnose early and to track disease severity progression automatically have gained increased importance. This thesis has explored the potential of building a multimodal automated method of measuring disease severity levels from patients that have been diagnosed with mild Alzheimer's disease.

Here, for the first time, new EEG-based diagnostic features have been explored for disease severity monitoring, as well as their fusion with MRI biomarkers. Four different machine learning algorithms were explored, as well as three methods to sift top-relevant features for the task at hand. Experimental results have shown that while the new EEG modulation spectral patch features are more relevant than conventional spectral power ones for severity monitoring, systems trained on MRI features tended to outperform EEG ones. Overall, multimodal EEG-MRI systems showed improved accuracy and lower variability across cross-validation runs, thus suggesting their importance for practical applications. Across different feature selection algorithms, the MRMR algorithm showed improved stability across different test setup runs and stood out among the other correlation-based candidates. Across machine learning algorithms, the non-linear regressors showed improved accuracy over linear ones.

The results presented in this thesis demonstrated the potential of the development of an automatic system able to predict the MMSE score of a patient with Alzheimer's disease in its mild stage,

utilizing biomarkers extracted from EEG and MRI. Moreover, the study may stimulate and encourage further research and development of such a system. Thus, clinicians, patients, and researchers could have access to an assistive automatic system to measure the cognitive assessment.

5.1 Study limitations

The results reported in this study were performed on a limited sample size of 89 participants. While repeated cross-validation and multiple trial runs allowed for data partitioning biases to be explored, the generalization capability of the models to unseen patients is still unknown. Moreover, to the best of the authors knowledge, no open-access dataset exists that includes EEG signals and MRI scans. The available data has a very narrow spread of MMSE scores, as shown in Figure 3.4. While this allows for models to be built for patients with mild levels of AD, it is still unknown how the models would behave as disease progresses into moderate and severe stages. As such, further analysis is still needed. Moreover, model performances showed an improvement when features that presented stability across the feature selection experiment runs were used. In that regard, more recent feature selection techniques should be explored in the search and identification of more stable selected features, in particular using wrapper methods in which the feature selection and the regressor are tightly coupled. Lastly, with the advances seen with deep neural network based analysis of [85] and MRI [58], deep multimodal architectures should be explored once more data is made available to the research community.

5.2 Future work

As for future research directions, work could concentrated on improving the performance of the automatic predictive models via better hyperparameters tuning, utilization and exploration of different fusion strategies, and the development of newer features. More importantly, evaluation of the obtained findings using a dataset comprised of patients with a much larger MMSE spread would be needed to explore if the models generalize well to more advanced stages of the disease. Moreover, with the availability of larger datasets, state-of-the-art deep neural network architectures could also be explored.

Bibliography

- [1] G. Aël Chetelat and J.-C. Baron. Early diagnosis of alzheimer's disease: contribution of structural neuroimaging. *Neuroimage*, 18(2):525–541, 2003.
- [2] A. Alberdi, A. Aztiria, and A. Basarab. On the early diagnosis of alzheimer's disease from multimodal signals: A survey. *Artificial Intelligence in Medicine*, 71:1–29, 2016.
- [3] A. Association. 2019 alzheimer's disease facts and figures. *Alzheimer's & Dementia*, 15(3):321–387, 2019.
- [4] P. K. Atrey, M. A. Hossain, A. El Saddik, and M. S. Kankanhalli. Multimodal fusion for multimedia analysis: a survey. *Multimedia systems*, 16(6):345–379, 2010.
- [5] A. Atri. The alzheimer's disease clinical spectrum: diagnosis and management. *Medical Clinics*, 103(2):263–293, 2019.
- [6] F. Bagherzadeh-Khiabani, A. Ramezankhani, F. Azizi, F. Hadaegh, E. W. Steyerberg, and D. Khalili. A tutorial on variable selection for clinical prediction models: feature selection methods in data mining could improve the results. *Journal of clinical epidemiology*, 71:76–85, 2016.
- [7] S. Basaia, F. Agosta, L. Wagner, E. Canu, G. Magnani, R. Santangelo, M. Filippi, A. D. N. Initiative, et al. Automated classification of alzheimer's disease and mild cognitive impairment using a single mri and deep neural networks. *NeuroImage: Clinical*, 21:101645, 2019.
- [8] S. Bauer, R. Wiest, L.-P. Nolte, and M. Reyes. A survey of mri-based medical image analysis for brain tumor studies. *Physics in Medicine & Biology*, 58(13):R97, 2013.

- [9] K. Bennys, G. Rondouin, C. Vergnes, and J. Touchon. Diagnostic value of quantitative eeg in alzheimer's disease. *Neurophysiologie Clinique/Clinical Neurophysiology*, 31(3):153–160, 2001.
- [10] C. Berka, D. J. Levendowski, M. N. Lumicao, A. Yau, G. Davis, V. T. Zivkovic, R. E. Olmstead, P. D. Tremoulet, and P. L. Craven. Eeg correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviation, space, and environmental medicine*, 78(5):B231–B244, 2007.
- [11] R. Black, B. Greenberg, J. M. Ryan, H. Posner, J. Seeburger, J. Amatniek, M. Resnick, R. Mohs, D. S. Miller, D. Saumier, et al. Scales as outcome measures for alzheimer's disease, 2009.
- [12] C. M. Bottino, C. C. Castro, R. L. Gomes, C. A. Buchpiguel, R. L. Marchetti, and M. R. L. Neto. Volumetric mri measurements can differentiate alzheimer's disease, mild cognitive impairment, and normal aging. *International Psychogeriatrics*, 14(1):59–72, 2002.
- [13] M. Bozzali, A. Falini, M. Franceschi, M. Cercignani, M. Zuffi, G. Scotti, G. Comi, and M. Filippi. White matter damage in alzheimer's disease assessed in vivo using diffusion tensor magnetic resonance imaging. *Journal of Neurology, Neurosurgery & Psychiatry*, 72(6):742–746, 2002.
- [14] R. Brookmeyer, E. Johnson, K. Ziegler-Graham, and H. M. Arrighi. Forecasting the global burden of alzheimer's disease. *Alzheimer's & dementia*, 3(3):186–191, 2007.
- [15] J. Cai, J. Luo, S. Wang, and S. Yang. Feature selection in machine learning: A new perspective. *Neurocomputing*, 300:70–79, 2018.
- [16] A. Caroli, G. Frisoni, A. D. N. Initiative, et al. The dynamics of alzheimer's disease biomarkers in the alzheimer's disease neuroimaging initiative cohort. *Neurobiology of aging*, 31(8):1263–1274, 2010.
- [17] D. A. Casey. Event-related potentials and the diagnosis of alzheimer's disease—the cognition™ system. *US Neurology*, 6(2):34–36, 2010.

- [18] R. Cassani, M. Estarellas, R. San-Martin, F. J. Fraga, and T. H. Falk. Systematic review on resting-state eeg for alzheimer’s disease diagnosis and progression assessment. *Disease markers*, 2018, 2018.
- [19] R. Cassani and T. Falk. Alzheimer’s disease diagnosis and severity level detection based on electroencephalography modulation spectral “patch” features. *IEEE Journal of Biomedical and Health Informatics*, 24(7):1982–1993, 2020.
- [20] R. Cassani and T. H. Falk. Automated alzheimer’s disease diagnosis using a low-density eeg layout and new features based on the power of modulation spectral “patches”. In *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, pages 1259–1263. IEEE, 2019.
- [21] R. Cassani, T. H. Falk, F. J. Fraga, M. Cecchi, D. K. Moore, and R. Anghinah. Towards automated electroencephalography-based alzheimer’s disease diagnosis using portable low-density devices. *Biomedical Signal Processing and Control*, 33:261–271, 2017.
- [22] R. Cassani, T. H. Falk, F. J. Fraga, P. A. Kanda, and R. Anghinah. The effects of automated artifact removal algorithms on electroencephalography-based alzheimer’s disease diagnosis. *Frontiers in aging neuroscience*, 6:55, 2014.
- [23] R. Cassani, T. H. Falk, F. J. Fraga, P. A. Kanda, and R. Anghinah. Towards automated eeg-based alzheimer’s disease diagnosis using relevance vector machines. In *5th ISSNIP-IEEE Biosignals and Biorobotics Conference (2014): Biosignals and Robotics for Better and Safer Living (BRC)*, pages 1–6. IEEE, 2014.
- [24] F. Castanedo. A review of data fusion techniques. *The Scientific World Journal*, 2013, 2013.
- [25] N. P. Castellanos and V. A. Makarov. Recovering eeg brain signals: artifact suppression with wavelet enhanced independent component analysis. *Journal of neuroscience methods*, 158(2):300–312, 2006.
- [26] M. Cecchi, D. K. Moore, C. H. Sadowsky, P. R. Solomon, P. M. Doraiswamy, C. D. Smith, G. A. Jicha, A. E. Budson, S. E. Arnold, and K. C. Fadem. A clinical trial to validate

- event-related potential markers of alzheimer's disease in outpatient settings. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 1(4):387–394, 2015.
- [27] N. S. Chok. *Pearson's versus Spearman's and Kendall's correlation coefficients for continuous data*. PhD thesis, University of Pittsburgh, 2010.
- [28] A. R. Clarke, R. J. Barry, R. McCarthy, and M. Selikowitz. Eeg analysis of children with attention-deficit/hyperactivity disorder and comorbid reading disabilities. *Journal of Learning Disabilities*, 35(3):276–285, 2002.
- [29] G. Constantine, K. Shan, S. D. Flamm, and M. U. Sivananthan. Role of mri in clinical cardiology. *The Lancet*, 363(9427):2162–2171, 2004.
- [30] J. Dauwels, F. Vialatte, and A. Cichocki. Diagnosis of alzheimer's disease from eeg signals: where are we standing? *Current Alzheimer Research*, 7(6):487–505, 2010.
- [31] J. Dauwels, F. Vialatte, T. Musha, and A. Cichocki. A comparative study of synchrony measures for the early diagnosis of alzheimer's disease based on eeg. *NeuroImage*, 49(1):668–693, 2010.
- [32] J. L. Devore. Probability and statistics for engineering and the sciences. 2008.
- [33] R. S. Doody, V. Pavlik, P. Massman, S. Rountree, E. Darby, and W. Chan. Predicting progression of alzheimer's disease. *Alzheimer's research & therapy*, 2(1):2, 2010.
- [34] B. Dumas, D. Lalanne, and S. Oviatt. Multimodal interfaces: A survey of principles, models and frameworks. In *Human machine interaction*, pages 3–26. Springer, 2009.
- [35] G. A. Edwards III, N. Gamez, G. Escobedo Jr, O. Calderon, and I. Moreno-Gonzalez. Modifiable risk factors for alzheimer's disease. *Frontiers in aging neuroscience*, 11, 2019.
- [36] S. F. Eskildsen, P. Coupé, V. S. Fonov, J. C. Pruessner, D. L. Collins, A. D. N. Initiative, et al. Structural imaging biomarkers of alzheimer's disease: predicting disease progression. *Neurobiology of aging*, 36:S23–S31, 2015.

- [37] F. Falahati, E. Westman, and A. Simmons. Multivariate data analysis and machine learning in alzheimer's disease with a focus on structural magnetic resonance imaging. *Journal of Alzheimer's disease*, 41(3):685–708, 2014.
- [38] T. H. Falk, F. J. Fraga, L. Trambaiolli, and R. Anghinah. Eeg amplitude modulation analysis for semi-automated diagnosis of alzheimer's disease. *EURASIP Journal on Advances in Signal Processing*, 2012(1):192, 2012.
- [39] F. Farina, D. Emek-Savaş, L. Rueda-Delgado, R. Boyle, H. Kiiski, G. Yener, and R. Whelan. A comparison of resting state eeg and structural mri for classifying alzheimer's disease and mild cognitive impairment. *bioRxiv*, 2019.
- [40] C. Fennema-Notestine, D. J. Hagler Jr, L. K. McEvoy, A. S. Fleisher, E. H. Wu, D. S. Karow, and A. M. Dale. Structural mri biomarkers for preclinical and mild alzheimer's disease. *Human brain mapping*, 30(10):3238–3253, 2009.
- [41] A. Filler. The history, development and impact of computed imaging in neurological diagnosis and neurosurgery: Ct, mri, and dti. *Nature proceedings*, pages 1–1, 2009.
- [42] B. Fischl. Freesurfer. *Neuroimage*, 62(2):774–781, 2012.
- [43] B. Fischl, D. H. Salat, E. Busa, M. Albert, M. Dieterich, C. Haselgrove, A. Van Der Kouwe, R. Killiany, D. Kennedy, S. Klaveness, et al. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*, 33(3):341–355, 2002.
- [44] B. Fischl, A. Van Der Kouwe, C. Destrieux, E. Halgren, F. Ségonne, D. H. Salat, E. Busa, L. J. Seidman, J. Goldstein, D. Kennedy, et al. Automatically parcellating the human cerebral cortex. *Cerebral cortex*, 14(1):11–22, 2004.
- [45] M. F. Folstein, S. E. Folstein, and P. R. McHugh. “mini-mental state”: a practical method for grading the cognitive state of patients for the clinician. *Journal of psychiatric research*, 12(3):189–198, 1975.
- [46] F. J. Fraga, T. H. Falk, P. A. Kanda, and R. Anghinah. Characterizing alzheimer's disease severity via resting-awake eeg amplitude modulation analysis. *PLoS One*, 8(8):e72240, 2013.

- [47] F. J. Fraga, T. H. Falk, L. R. Trambaiolli, E. F. Oliveira, W. H. Pinaya, P. A. Kanda, and R. Anghinah. Towards an eeg-based biomarker for alzheimer’s disease: improving amplitude modulation analysis features. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 1207–1211. IEEE, 2013.
- [48] G. B. Frisoni, N. C. Fox, C. R. Jack, P. Scheltens, and P. M. Thompson. The clinical use of structural mri in alzheimer disease. *Nature Reviews Neurology*, 6(2):67–77, 2010.
- [49] M. M. Grigg, M. A. Kelly, G. G. Celesia, M. W. Ghobrial, and E. R. Ross. Electroencephalographic activity after brain death. *Archives of neurology*, 44(9):948–954, 1987.
- [50] A. L. Gross, D. M. Mungas, J.-M. S. Leoutsakos, M. S. Albert, R. N. Jones, A. D. N. Initiative, et al. Alzheimer’s disease severity, objectively determined and measured. *Alzheimer’s & Dementia: Diagnosis, Assessment & Disease Monitoring*, 4:159–168, 2016.
- [51] Y. Gupta, K. H. Lee, K. Y. Choi, J. J. Lee, B. C. Kim, G. R. Kwon, N. R. C. for Dementia, and A. D. N. Initiative. Early diagnosis of alzheimer’s disease using combined features from voxel-based morphometry and cortical, subcortical, and hippocampus regions of mri t1 brain images. *PLoS one*, 14(10):e0222446, 2019.
- [52] H. Hampel, R. Frank, K. Broich, S. J. Teipel, R. G. Katz, J. Hardy, K. Herholz, A. L. Bokde, F. Jessen, Y. C. Hoessler, et al. Biomarkers for alzheimer’s disease: academic, industry and regulatory perspectives. *Nature reviews Drug discovery*, 9(7):560–574, 2010.
- [53] L. E. Hebert, J. Weuve, P. A. Scherr, and D. A. Evans. Alzheimer disease in the united states (2010–2050) estimated using the 2010 census. *Neurology*, 80(19):1778–1783, 2013.
- [54] H. Hippius and G. Neundörfer. The discovery of alzheimer’s disease. *Dialogues in clinical neuroscience*, 5(1):101, 2003.
- [55] B. M. Jedynak, A. Lang, B. Liu, E. Katz, Y. Zhang, B. T. Wyman, D. Raunig, C. P. Jedynak, B. Caffo, J. L. Prince, et al. A computational neurodegenerative disease progression score: method and results with the alzheimer’s disease neuroimaging initiative cohort. *Neuroimage*, 63(3):1478–1486, 2012.

- [56] B. Jelles, P. Scheltens, W. Van der Flier, E. Jonkman, F. L. da Silva, and C. Stam. Global dynamical analysis of the eeg in alzheimer's disease: frequency-specific changes of functional interactions. *Clinical neurophysiology*, 119(4):837–841, 2008.
- [57] J. Jeong. Eeg dynamics in patients with alzheimer's disease. *Clinical neurophysiology*, 115(7):1490–1505, 2004.
- [58] T. Jo, K. Nho, and A. J. Saykin. Deep learning in alzheimer's disease: diagnostic classification and prognostic prediction using neuroimaging data. *Frontiers in aging neuroscience*, 11:220, 2019.
- [59] A. Jović, K. Brkić, and N. Bogunović. A review of feature selection methods with applications. In *2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, pages 1200–1205. IEEE, 2015.
- [60] N. Kannathal, M. L. Choo, U. R. Acharya, and P. Sadasivan. Entropies for detection of epilepsy in eeg. *Computer methods and programs in biomedicine*, 80(3):187–194, 2005.
- [61] E. G. Kehoe, J. P. McNulty, P. G. Mullins, and A. L. Bokde. Advances in mri biomarkers for the diagnosis of alzheimer's disease. *Biomarkers in medicine*, 8(9):1151–1169, 2014.
- [62] U. M. Khaire and R. Dhanalakshmi. Stability of feature selection algorithm: A review. *Journal of King Saud University-Computer and Information Sciences*, 2019.
- [63] H. Koponen, J. Partanen, A. Pääkkönen, E. Mattila, and P. Riekkinen. Eeg spectral analysis in delirium. *Journal of Neurology, Neurosurgery & Psychiatry*, 52(8):980–985, 1989.
- [64] J. W. Kowalski, M. Gawel, A. Pfeffer, and M. Barcikowska. The diagnostic value of eeg in alzheimer disease: correlation with the severity of mental impairment. *Journal of clinical neurophysiology*, 18(6):570–575, 2001.
- [65] M. A. Kuskowski, J. A. Mortimer, G. K. Morley, S. M. Malone, and A. J. Okaya. Rate of cognitive decline in alzheimer's disease is associated with eeg alpha power. *Biological psychiatry*, 33(8-9):659–662, 1993.

- [66] Y. Lakmache, M. Lassonde, S. Gauthier, J.-Y. Frigon, and F. Lepore. Interhemispheric disconnection syndrome in alzheimer's disease. *Proceedings of the National Academy of Sciences*, 95(15):9042–9046, 1998.
- [67] S. Leandrou, S. Petroudi, P. A. Kyriacou, C. C. Reyes-Aldasoro, and C. S. Pattichis. Quantitative mri brain studies in mild cognitive impairment and alzheimer's disease: a methodological review. *IEEE reviews in biomedical engineering*, 11:97–111, 2018.
- [68] A. Lebedev, E. Westman, G. Van Westen, M. Kramberger, A. Lundervold, D. Aarsland, H. Soininen, I. Kloszewska, P. Mecocci, M. Tsolaki, et al. Random forest ensembles for detection and prediction of alzheimer's disease with a good between-cohort robustness. *NeuroImage: Clinical*, 6:115–125, 2014.
- [69] J. Li, K. Cheng, S. Wang, F. Morstatter, R. P. Trevino, J. Tang, and H. Liu. Feature selection: A data perspective. *ACM Computing Surveys (CSUR)*, 50(6):1–45, 2017.
- [70] X. Long, L. Zhang, W. Liao, C. Jiang, B. Qiu, and A. D. N. Initiative. Distinct laterality alterations distinguish mild cognitive impairment and alzheimer's disease from healthy aging: statistical parametric mapping with high resolution mri. *Human brain mapping*, 34(12):3400–3410, 2013.
- [71] N. Mammone, F. La Foresta, and F. C. Morabito. Automatic artifact rejection from multi-channel scalp eeg by wavelet ica. *IEEE Sensors Journal*, 12(3):533–542, 2011.
- [72] U. G. Mangai, S. Samanta, S. Das, and P. R. Chowdhury. A survey of decision fusion and feature fusion strategies for pattern classification. *IETE Technical review*, 27(4):293–307, 2010.
- [73] A. Martínez-Torteya, V. Treviño, and J. G. Tamez-Peña. Improved diagnostic multimodal biomarkers for alzheimer's disease and mild cognitive impairment. *BioMed research international*, 2015, 2015.
- [74] R. Mayeux and N. Schupf. Blood-based biomarkers for alzheimer's disease: plasma $\alpha\beta40$ and $\alpha\beta42$, and genetic variants. *Neurobiology of aging*, 32:S10–S19, 2011.

- [75] W. J. McGeown, R. Cassani, T. H. Falk, M. Cecchi, and K. Fadem. P2-364: Neuroanatomical and neuropsychological correlates of resting state eeg diagnostic features in patients with alzheimer's disease. *Alzheimer's & Dementia*, 15(7S_Part_14):P739–P740, 2019.
- [76] Z. S. Nasreddine, N. A. Phillips, V. Bédirian, S. Charbonneau, V. Whitehead, I. Collin, J. L. Cummings, and H. Chertkow. The montreal cognitive assessment, moca: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4):695–699, 2005.
- [77] A. Nordberg, J. O. Rinne, A. Kadir, and B. Långström. The use of pet in alzheimer disease. *Nature Reviews Neurology*, 6(2):78, 2010.
- [78] M. R. Nuwer, D. A. Hovda, L. M. Schrader, and P. M. Vespa. Routine and quantitative eeg in mild traumatic brain injury. *Clinical Neurophysiology*, 116(9):2001–2025, 2005.
- [79] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al. Scikit-learn: Machine learning in python. *the Journal of machine Learning research*, 12:2825–2830, 2011.
- [80] H. Peng, F. Long, and C. Ding. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on pattern analysis and machine intelligence*, 27(8):1226–1238, 2005.
- [81] C. Plant, S. J. Teipel, A. Oswald, C. Böhm, T. Meindl, J. Mourao-Miranda, A. W. Bokde, H. Hampel, and M. Ewers. Automated detection of brain atrophy patterns based on mri for the prediction of alzheimer's disease. *Neuroimage*, 50(1):162–174, 2010.
- [82] Z. Qin, A. T. Wang, C. Zhang, and S. Zhang. Cost-sensitive classification with k-nearest neighbors. In *International Conference on Knowledge Science, Engineering and Management*, pages 112–131. Springer, 2013.
- [83] B. C. Riedel, P. M. Thompson, and R. D. Brinton. Age, apoe and sex: triad of risk of alzheimer's disease. *The Journal of steroid biochemistry and molecular biology*, 160:134–147, 2016.

- [84] R. Rodríguez-Perez, M. Vogt, and J. Bajorath. Support vector machine classification and regression prioritize different structural features for binary compound activity and potency value prediction. *ACS omega*, 2(10):6371–6379, 2017.
- [85] Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. H. Falk, and J. Faubert. Deep learning-based electroencephalography analysis: a systematic review. *Journal of neural engineering*, 16(5):051001, 2019.
- [86] B. Saletu, P. Anderer, G. Saletu-Zyhlarz, and R. Pascual-Marqui. Eeg topography and tomography in diagnosis and treatment of mental disorders: evidence for a key-lock principle. *Methods and findings in experimental and clinical pharmacology*, 24:97–106, 2002.
- [87] M. C. Salinsky, B. S. Oken, D. Storzbach, and C. B. Dodrill. Assessment of cns effects of antiepileptic drugs by using quantitative eeg measures. *Epilepsia*, 44(8):1042–1050, 2003.
- [88] S. B. Sando, S. Melquist, A. Cannon, M. Hutton, O. Sletvold, I. Saltvedt, L. R. White, S. Lydersen, and J. Aasly. Risk-reducing effect of education in alzheimer’s disease. *International Journal of Geriatric Psychiatry: A journal of the psychiatry of late life and allied sciences*, 23(11):1156–1162, 2008.
- [89] M. J. Sands and A. Levitin. Basics of magnetic resonance imaging. In *Seminars in vascular surgery*, volume 17, pages 66–82. Elsevier, 2004.
- [90] S. Sanei and J. A. Chambers. *EEG signal processing*. John Wiley & Sons, 2013.
- [91] M. E. Shenton, H. Hamoda, J. Schneiderman, S. Bouix, O. Pasternak, Y. Rathi, M.-A. Vu, M. P. Purohit, K. Helmer, I. Koerte, et al. A review of magnetic resonance imaging and diffusion tensor imaging findings in mild traumatic brain injury. *Brain imaging and behavior*, 6(2):137–192, 2012.
- [92] H. W. Shipton. Eeg analysis: A history and a prospectus. *Annual review of biophysics and bioengineering*, 4(1):1–13, 1975.
- [93] M. V. F. Silva, C. d. M. G. Loures, L. C. V. Alves, L. C. de Souza, K. B. G. Borges, and M. das Graças Carvalho. Alzheimer’s disease: risk factors and potentially protective measures. *Journal of biomedical science*, 26(1):33, 2019.

- [94] A. J. Smola and B. Schölkopf. A tutorial on support vector regression. *Statistics and computing*, 14(3):199–222, 2004.
- [95] K. Strimbu and J. A. Tavel. What are biomarkers? *Current Opinion in HIV and AIDS*, 5(6):463, 2010.
- [96] R. E. Tanzi. The genetics of alzheimer disease. *Cold Spring Harbor perspectives in medicine*, 2(10):a006296, 2012.
- [97] K. A. Townsend, G. Wollstein, and J. S. Schuman. Clinical application of mri in ophthalmology. *NMR in Biomedicine: An International Journal Devoted to the Development and Application of Magnetic Resonance In vivo*, 21(9):997–1002, 2008.
- [98] B. Türkbey, D. Thomasson, Y. Pang, M. Bernardo, and P. L. Choyke. The role of dynamic contrast-enhanced mri in cancer diagnosis and treatment. *Diagnostic and interventional radiology (Ankara, Turkey)*, 16(3):186, 2010.
- [99] C. Van Cauwenberghe, C. Van Broeckhoven, and K. Sleegers. The genetic landscape of alzheimer disease: clinical implications and perspectives. *Genetics in Medicine*, 18(5):421–430, 2016.
- [100] K. Viau. Poster presentations at the 10th canadian conference on dementia (ccd) québec city, october 3, 2019.
- [101] K. Wild, D. Howieson, F. Webbe, A. Seelye, and J. Kaye. Status of computerized cognitive testing in aging: a systematic review. *Alzheimer's & Dementia*, 4(6):428–437, 2008.
- [102] D. Zhang, Y. Wang, L. Zhou, H. Yuan, D. Shen, A. D. N. Initiative, et al. Multimodal classification of alzheimer’s disease and mild cognitive impairment. *Neuroimage*, 55(3):856–867, 2011.