



HAL
open science

Multi-scale and multimodal imaging biomarkers for the early detection of Alzheimer's disease

Kilian Hett

► **To cite this version:**

Kilian Hett. Multi-scale and multimodal imaging biomarkers for the early detection of Alzheimer's disease. Image Processing [eess.IV]. Université de Bordeaux, 2019. English. NNT : 2019BORD0011 . tel-02102247

HAL Id: tel-02102247

<https://theses.hal.science/tel-02102247>

Submitted on 17 Apr 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

présentée à

L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE DE MATHÉMATIQUES ET D'INFORMATIQUE

par

Kilian Hett

POUR OBTENIR LE GRADE DE

DOCTEUR

SPÉCIALITÉ : Informatique

**Nouveaux biomarqueurs multi-échelles et
multi-modaux pour le diagnostic précoce de la
maladie d'Alzheimer**

**Multi-scale and multimodal imaging biomarkers
for the early detection of Alzheimer's disease**

Date de soutenance : 25 Janvier 2019**Devant la commission d'examen composée de :**

Jean-Philippe Su	DOMENGER	Directeur de recherche	University of Bordeaux	Président
Olivier	RUAN	Professeur des universités	Université de Rouen	Rapporteuse
Christian	COLLIOT	Directeur de recherche	CNRS	Rapporteur
Mallar	BARILLOT	Directeur de recherche	CNRS	Rapporteur
Pierrick	CHAKRAVARTY	Associate professor	McGill University	Examinateur
Vinh-Thong	COUPÉ	Chargé de recherche	CNRS	Directeur de thèse
	TA	Maître de conférences	INP-Bordeaux	Co-superviseur

Titre

Nouveaux biomarqueurs multi-échelles et multi-modaux pour le diagnostic précoce de la maladie d'Alzheimer

Résumé

La maladie d'Alzheimer est la première cause de démence chez les personnes âgées. Cette maladie est caractérisée par un déclin irréversible des fonctions cognitives. Les patients atteints par la maladie d'Alzheimer ont de sévères pertes de mémoire et ont de grandes difficultés à apprendre de nouvelles informations ce qui pose de gros problèmes dans leur vie quotidienne. À ce jour, cette maladie est diagnostiquée après que d'importantes altérations des structures du cerveaux apparaissent. De plus, aucune thérapie existe permettant de faire reculer ou de stopper la maladie. Le développement de nouvelles méthodes permettant la détection précoce de cette maladie est ainsi nécessaire. En effet, une détection précoce permettrait une meilleure prise en charge des patients atteints de cette maladie ainsi qu'une accélération de la recherche thérapeutique. Nos travaux de recherche portent sur l'utilisation de l'imagerie médicale, avec notamment l'imagerie par résonance magnétique (IRM) qui a démontrée ces dernières années son potentiel pour améliorer la détection et la prédiction de la maladie d'Alzheimer. Afin d'exploiter pleinement ce type d'imagerie, de nombreuses méthodes ont été proposées récemment. Au cours de nos recherches, nous nous sommes intéressés à un type de méthode en particulier qui est basé sur la correspondance de patches dans de grandes bibliothèques d'images. Nous avons étudié ces méthodes à diverses échelles anatomiques (c'est à dire, cerveaux entier, hippocampe, sous-champs de l'hippocampe) avec diverses modalités d'IRM (par exemple, IRM anatomique et imagerie de diffusion). Nous avons amélioré les performances de détection dans les stades les plus précoces avec l'imagerie par diffusion. Nous avons aussi proposé un nouveau schéma de fusion pour combiner IRM anatomique et imagerie de diffusion. De plus, nous avons montré que la correspondance de patches était améliorée par l'utilisation de filtres dérivatifs. Enfin, nous avons proposé une méthode par graphe permettant de combiner les informations de similarité inter-sujet avec les informations apportées par la variabilité intra-sujet. Les résultats des expériences menées dans cette thèse ont montrées une amélioration des performances de diagnostique et de pronostique de la maladie d'Alzheimer comparé aux méthodes de l'état de l'art.

Mots-clés

Imagerie médicale, Maladie d'Alzheimer, Multi-modalité, Multi-échelle, Analyse du cerveaux, Méthodes non-local

Title

Multi-scale and multimodal imaging biomarkers for the early detection of Alzheimer's disease

Abstract

Alzheimer's disease (AD) is the most common dementia leading to a neurodegenerative process and causing mental dysfunctions. According to the world health organization, the number of patients having AD will double in 20 years. Neuroimaging studies performed on AD patients revealed that structural brain alterations are advanced when the diagnosis is established. Indeed, the clinical symptoms of AD are preceded by brain changes. This stresses the need to develop new biomarkers to detect the first stages of the disease. The development of such biomarkers can make easier the design of clinical trials and therefore accelerate the development of new therapies. Over the past decades, the improvement of magnetic resonance imaging (MRI) has led to the development of new imaging biomarkers. Such biomarkers demonstrated their relevance for computer-aided diagnosis but have shown limited performances for AD prognosis. Recently, advanced biomarkers were proposed to improve computer-aided prognosis. Among them, patch-based grading methods demonstrated competitive results to detect subtle modifications at the earliest stages of AD. Such methods have shown their ability to predict AD several years before the conversion to dementia. For these reasons, we have had a particular interest in patch-based grading methods. First, we studied patch-based grading methods for different anatomical scales (*i.e.*, whole brain, hippocampus, and hippocampal subfields). We adapted patch-based grading method to different MRI modalities (*i.e.*, anatomical MRI and diffusion-weighted MRI) and developed an adaptive fusion scheme. Then, we showed that patch comparisons are improved with the use of multi-directional derivative features. Finally, we proposed a new method based on a graph modeling that enables to combine information from inter-subjects' similarities and intra-subjects' variability. The conducted experiments demonstrate that our proposed method enable an improvement of AD detection and prediction.

Keywords

Medical imaging, Alzheimer's disease, Multimodality, Multi-scale, Brain analysis, Patch-based methods

Unité de recherche

Université de Bordeaux, LaBRI, UMR 5800, PICTURA, F-33400 Talence, France
CNRS, LaBRI, UMR 5800, PICTURA, F-33400 Talence, France

Remerciement

Je souhaiterai en premier lieu remercier Su Ruan, Olivier Colliot et Christian Barillot pour avoir accepté de rapporter ma thèse. Merci aussi à Mallar Chakravarty pour avoir accepté de faire partie de mon jury et avoir fait le déplacement depuis le Canada. Enfin, merci à Jean-Philippe Domenger pour avoir été président de mon jury.

Durant ces trois années, j'ai eu la chance d'être encadré par Pierrick Coupé et Vinh-Thong Ta. Je tiens particulièrement à les remercier pour leur exigence, leur disponibilité, et leur patience qui a été à toutes épreuves lors de la relecture de mes papiers. Merci à Pierrick, pour son investissement, son aide et les discussions à propos de mes travaux qui à chaque fois étaient une illumination scientifique pour moi. Merci à Vinh, en particulier pour ses longues heures de discussion, parfois nocturnes, concernant les papiers sur lesquels je travaillais (le plancher qui grince à trois heures du matin ne me fait plus peur). Travailler avec eux m'a permis de progresser sur tant d'aspects (méthodologique, scientifique, rédactionnel, organisationnel, ...). Je leur dois énormément, et leur en suis infiniment reconnaissant. J'espère avoir l'opportunité de pouvoir continuer à collaborer avec eux dans le futur.

Souvent, je me pose cette question: "Qu'aurait été ces années à Bordeaux sans mes amis: Rémi, Florent, Joachim, Christelle, Cyril, Lucile, Léa, Juliette, Pierre, Gaël (mieux connu à Bordeaux sous le nom de High-kick Man), Laure, Aurélie, Noémie, Estelle?". Avec certitude, je peux aujourd'hui y répondre. Ça aurait été vraiment moins sympa. Un merci particulier à Rémi pour son enseignement quasi religieux du ping-pong (le top-spin, la poussette ainsi que le service "ambulance" n'ont plus aucun secret pour moi). Enseignement qui m'a permis de souvent éviter le port de cette fameuse casquette rouge, merci à Florent pour son fair-play et son respect de la parole donnée (clin d'oeil). Un immense merci, général cette fois, à toutes les personnes que j'ai citées pour avoir partagé avec moi cette aventure Bordelaise et pour avoir rendu agréable ces années de thèse.

Je tiens ensuite à remercier ma famille. Merci à mes parents pour leur soutien inconditionnel, et pour leur si nombreux enseignements: dont la patience, la persévérance et l'éthique. Qualités qui m'ont été grandement nécessaires pour venir à bout de ces années de thèse. Merci à mon frère et ma soeur pour leur présence. Vous avez donné du sens et du relief au mot "famille". Pour tout ce qu'ils ont fait pour moi, je leur suis éternellement reconnaissant.

Enfin, je souhaite remercier ma petite amie, Kayla, pour sa douceur, sa patience et son affection. Je suis particulièrement reconnaissant pour sa présence, et ceux malgré l'océan Atlantique qui nous a séparé pendant deux ans.

Résumé en français

Introduction

Maladie d'Alzheimer

La maladie d'Alzheimer (AD) est la démence la plus répandue chez les personnes âgées. La prévalence de cette maladie est d'environ 1% chez les personnes de 60 ans et augmente pour atteindre environ 40% des personnes de 90 ans. De plus, selon l'organisation mondiale de la santé, avec le vieillissement constant de la population, le nombre de patients atteints d'Alzheimer doublera dans les 20 prochaines années. Par conséquent, la maladie d'Alzheimer va représenter un coût financier important pour la société dans la prochaine décennies. Les patients atteints de cette démence sont touchés par un processus neurodégénératif irréversible qui cause des dysfonctionnements cognitifs tels que la perte de mémoire à long terme, des altérations du langage, une désorientation, un changement de personnalité, et qui finalement entraînent la mort. Cette maladie est caractérisée par une accumulation de plaques de bêta-amyloïde et d'enchevêtrements neurofibrillaires composé de neurofibrilles amyloïdes tau associé à une perte des synapses et neurones. À ce jour, aucun traitement connu n'a été capable d'arrêter ou de ralentir la progression de cette maladie. De plus, la neuroimagerie a révélé que des modifications du cerveau se produisent des décennies avant le diagnostic est établi. Enfin, l'évolution de la charge pathologique n'est pas corrélée linéairement à la fonction cognitive. Ainsi, lorsque le diagnostic de la maladie d'Alzheimer la maladie est établie, la charge pathologique est déjà élevée.

Une phase prodromique de la maladie d'Alzheimer est le déficit cognitif léger (MCI). Les symptômes cliniques des patients atteints de MCI sont légers même si la diminution des capacités cognitives est mesurable. Le syndrome MCI peut être provoqué par des sources de facteur hétérogènes telles que, dans la plupart des cas, une maladie cérébrovasculaire, une démence à corps de Lewy, une démence fronto-temporale, aphasia primaire progressive. En outre, des études antérieures ont suggéré qu'environ 12% des sujets atteints par le syndrome MCI progresse vers la maladie d'Alzheimer au cours des quatre années suivant l'apparition des premiers symptômes. Cependant, bien que les sujets MCI présentent un risque élevé de développer la maladie d'Alzheimer, les sujets souffrant de ce syndrome peuvent rester stables (c'est à dire, ne pas convertir vers Alzheimer ou en une autre démence). Ce groupe de patients est appelé MCI stable à l'opposé de les patients qui développent la maladie d'Alzheimer dans les années qui suivent le diagnostic de symptômes de

MCI. Ce dernier groupe de patient est appelé MCI progressive (pMCI). Le sujet souffrant de MCI peut également se convertir en une autre démence, voire même revenir à un état cognitif normal. La prédiction précoce des sujets souffrant de symptômes de MCI qui convertiront vers la maladie d'Alzheimer pourrait améliorer l'efficacité des futurs traitements en réduisant les altérations cérébrales avant le début du traitement. De plus, la prédiction de la conversion pourrait également accélérer le développement de nouvelles thérapies en rendant la sélection du sujet plus précise, ce qui réduirait le coût des essais cliniques et permettrait des études cliniques plus précises.

Imagerie médicale pour la détection

Avec l'amélioration récente des techniques d'imagerie médicale qui fournissent des outils d'imagerie *in vivo* puissants et non invasifs, de nouveaux biomarqueurs d'imagerie ont été proposés. Ces biomarqueurs peuvent être regroupés en fonction du type d'informations physiques qu'ils capturent. Les biomarqueurs basés sur l'imagerie par résonance magnétique (IRM) capturent principalement les altérations structurelles des structures cérébrales telles que l'atrophie de l'hippocampe et hypertrophie des ventricules latéraux. Les biomarqueurs nucléaires utilisant l'imagerie nucléaire tels que la tomographie par émission de positrons (TEP). Cette technique d'imagerie détecte la présence de protéines spécifiques telles que les dépôts anormaux de bêta-amyloïdes dans l'hippocampe. Enfin, les biomarqueurs fonctionnels tirant avantage de l'imagerie fonctionnelle ont été développés pour décrire les altérations des fonction cognitive.

Les biomarqueurs utilisant l'IRM sont des techniques *in vivo* qui proposent plusieurs avantages. Les techniques employé sont non invasives, elles fournissent des détails structurels et fonctionnels du cerveau presque immédiatement, ils sont sensibles et spécifiques. De plus, les biomarqueurs d'imagerie fournissent un diagnostic qui est reproductible. Par conséquent, même si ce type de biomarqueur est coûteux nécessite du personnel expérimenté, le développement de nouveaux biomarqueurs d'imagerie a été très intense ces dernières années.

En effet, il a été démontré que les méthodes d'apprentissage automatique ont le potentiel pour aider à identifier les patients atteints par la maladie d'Alzheimer en découvrant des schémas discriminants dans les données de neuroimagerie. À ce jour, les biomarqueurs basés sur l'IRM ont atteint de bonnes performances pour le diagnostic, mais sont encore limitées pour le pronostic de la maladie. Par conséquent, dans cette thèse, nous nous sommes concentré sur le développement de nouveaux biomarqueurs d'imagerie utilisant l'IRM dans le but d'améliorer la détections des altérations structurelles.

Méthodes basées sur les patches

L'utilisation de méthodes basées sur les patches pour la détection de la maladie d'Alzheimer a été initialement proposée. posé par [Coupé et al. \(2012b\)](#) avec le cadre de classement basé sur les patches (PBG). L'idée principale de cette méthode basée sur des exemples est d'utiliser la capacité des techniques basée sur les patches pour capturer des altérations subtiles du signal. L'idée est de propager les informations de motif local intégrées par une bibliothèque de modèles à l'image en cours d'analyse. À l'origine, cette méthode a été proposée pour capturer les dégradations anatomiques causées par AD. Afin de déterminer le statut pathologique du sujet en cours, la méthode PBG estime à chaque voxel la gravité des altérations structurelles par une mesure de similarité non locale. La méthode SNIPE augmente les performances de classification des méthodes basées sur une analyse de la structure de l'hippocampe. En effet, la méthode PBG a obtenu une précision de 88% pour la détection de la maladie et une précision de 71% pour sa prédiction. Cette méthode a dépassé les approches précédentes basées sur l'analyse de l'hippocampe.

Motivations

Dans cette thèse, nous nous sommes concentrés sur le développement de nouveaux biomarqueurs efficaces pour la détection précoce de la maladie d'Alzheimer. En particulier, nous nous sommes principalement intéressés à l'amélioration des méthodes de classement par patch. En effet, l'introduction de méthodes basées sur les patches a permis de mieux estimer les modifications structurelles de structures spécifiques telles que l'hippocampe. Cependant, certains éléments tendent à montrer qu'une meilleure analyse de l'hippocampe pourrait augmenter les performances pour la détection précoce de la maladie d'Alzheimer.

En effet, d'une part, une grande partie des données proviennent de différentes campagnes d'acquisition et ont été obtenues dans différents centres de neuroimagerie. Par conséquent, même si un protocole d'acquisition a été défini, les images sont acquises avec différents dispositifs d'imagerie. Cela se traduit par une variabilité des intensités des images qui pourraient ne pas être corrigées par les étapes de prétraitement. Ainsi, il semble essentiel de développer une méthode robuste à la variabilité d'acquisition. En outre, l'amélioration de la comparaison des patches en mettant en évidence le signal discriminant dans les patches pourrait conduire à une amélioration des performances pour la détection précoce de la maladie. En effet, seulement une partie du signal contenu dans un patch peut représenter toutes l'information discriminante. Par exemple, les signaux de bord et de texture pourraient être discriminants lorsque l'autre partie du signal pourrait ne pas être informative et simplement ajouter du bruit au sein de chaque groupe pathologique, car étant peut être corrélée à d'autres facteurs.

D'autre part, l'IRM structurelle (IRM-s) n'est pas la seule modalité d'IRM utilisable dans le cadre des méthodes PBG. Par exemple, l'IRM de diffusion (IRM-d) peut

également être intégrée dans les méthodes PBG. Les études basées sur la modalité IRM-d capturent généralement les altérations des axones en analysant les modifications de la substance blanche par des approches de tractographie. Cependant, l'IRM-d peut également être utilisée pour capturer les modifications des microstructures apparaissant dans les structures composées de matière grise. L'application de la méthode PBG sur cette modalité pourrait améliorer les performances de détection précoce car de telles altérations sont considérées comme se produisant avant modification structurelle. Cependant, une limitation de la d-IRM provenant de sa résolution native, qui est souvent inférieure à celle de la IRM-s, réduit sa capacité à fournir de bons biomarqueurs pour la détection AD. Néanmoins, les méthodes récentes de super-résolution ont donné de bons résultats et ont résolu des problèmes provoqués par la basse résolution des images dans de nombreux problèmes de reconnaissance de formes.

En outre, l'hippocampe est une structure hétérogène. En effet, l'hippocampe est composé de différents sous-champs ayant des caractéristiques distinctes. De plus, au cours de la dernière décennie, Des études post mortem et chez l'animal ont montré que la MA ne touchait pas les sous-champs de l'hippocampe. En effet, ces études ont montré que certains sous-champs souffrent d'une atrophie plus forte à la fin de l'AD que d'autres. Par conséquent, une analyse fine des modifications structurelles de l'hippocampe à l'échelle des sous-champs pourrait également permettre d'améliorer la détection précoce de la maladie.

Enfin, bien que les méthodes de l'état de l'art pour la détection précoce de la maladie d'Alzheimer aient été principalement axées sur une extraction robuste des similarités inter-sujets dans des régions d'intérêt spécifiques ou différents l'échelle d'analyse (c'est à dire, voxel, patch, etc), certains travaux ont mis au point des modèles qui rendent compte de la variabilité des altérations au sein du même sujets. Ces méthodes ont démontré qu'une représentation efficace de la topologie des modifications, en modélisant la relation des modifications structurelles entre les différentes structures du cerveau, peut aider à améliorer la détection et la prédiction de la maladie d'Alzheimer. En effet, il semble que les relations des altérations structurelles entre les structures cérébrales forment un schéma spécifique qui fournit des informations pertinentes sur l'évolution de la démence. Par conséquent, la modélisation de ce modèle pourrait améliorer la détection précoce de la maladie.

En conséquence, pour répondre aux questions soulevées par ces éléments nous avons exploré deux grands axes de recherche. Le développement de biomarqueur hippocampal avancé et le développement d'un modèle de la signature des altérations structurelle provoqué par la maladie avec une approche basée graphe.

Biomarqueur avancé basée sur l’hippocampe

Dans la deuxième partie de cette thèse, nous avons développé des biomarqueurs avancés. Nous avons d’abord mis au point de nouvelles méthodes que nous avons validées au sein de l’hippocampe. Nous avons proposé une méthode basée sur les approches PBG en y intégrant des filtres de texture qui permet d’obtenir des performances de pointe pour la détection AD. Cette première contribution démontre que l’intensité de l’IRM n’est pas la meilleure entrée et qu’un filtre dérivatif multidirectionnelle, tel que le filtre de Gabor, permet de mettre en évidence des signaux informatifs. Ensuite, nous avons développé une méthode PBG multimodale combinant IRM-s et MRI-d en utilisant un schéma de fusion adaptative. Les expériences menées dans ce deuxième travail ont montré que, bien que l’IRM-s soit toujours une bonne entrée pour la détection de la maladie d’Alzheimer, les changements micro-structuraux détectés avec la modalité IRM-d et capturés avec les méthodes PBG permettent une meilleure prédiction de la progression de la maladie. De plus, notre méthode multimodale a démontré sa capacité à calculer un biomarqueur hippocampique robuste qui a obtenu les meilleurs résultats pour la MA. détection et prédiction.

Deuxièmement, nous avons étudié l’hippocampe à une échelle anatomique niveau plus fine. En effet, au lieu de considérer l’hippocampe avec une approche globale, nous avons étudié l’efficacité des sous-champs de l’hippocampe. Nous avons étudié les altérations des sous-champs de l’hippocampe avec notre nouvelle méthode MPBG. Nos expériences montrent que le subiculum souffre des changements les plus significatifs aux premiers stades de la maladie d’Alzheimer. En effet, les méthodes MPBG appliquées au subiculum permettent d’augmenter la performances de prédiction par rapport à son application dans l’ensemble de la structure de l’hippocampe. Ces résultats confirment que les sous-champs les plus discriminants de l’hippocampe permettent d’obtenir des biomarqueurs plus efficaces.

Enfin, bien que les contributions présentées dans cette partie présentent des résultats de pointe en matière de détection, la prédiction de la conversion des sujets en la maladie d’Alzheimer reste limitée. En effet, les améliorations apportées avec les nouvelles améliorations proposées, par l’utilisation de modalités multiples ou d’échelle d’analyse plus précis n’ont pas permis d’augmenter de manière significative les performances de classification pour la prédiction AD. Notre hypothèse principale est que même si notre méthode permet de capturer des modifications structurelles subtiles, une meilleure modélisation des relation des altérations apparaissant entre les structures considérées pourrait fournir des informations utiles. Par conséquent, dans la suite de cette thèse, nous développerons une nouvelle méthode basée sur une modélisation par graphe qui combine les informations de similarité inter-sujets et la variabilité intra-sujets.

Modélisation de la signature de la maladie d’Alzheimer

Dans cette partie, nous avons développé une nouvelle méthode pour mieux modéliser la signature de la maladie d’Alzheimer. Le modèle de la maladie d’Alzheimer développé pour cette méthode intègre deux types d’informations différentes. La première est une information basée sur les similitudes inter-sujets extraites avec une méthode PBG. La seconde intègre une représentation de la variabilité intra-sujets des altérations capturées avec une modélisation basée sur des graphes. Nos expériences ont démontré que notre modélisation combinant ces deux types d’informations permet une prédiction plus efficace de la conversion des sujets à la maladie d’Alzheimer. De plus, notre graphe des altérations des structures est générique et peut être appliqué avec diverses représentations de structures cérébrales. En effet, nous avons appliqué notre méthode basée graphe à différentes échelles anatomiques. Dans les deux cas, les expériences ont montré une augmentation des performances de prédiction de la maladie d’Alzheimer. Enfin, nous avons proposé une méthode combinant plusieurs échelles anatomiques du cerveau. Nous avons validé notre méthode pour la combinaison des structures cérébrales et des sous-champs de l’hippocampe. Les expériences réalisées dans cette partie ont montré des résultats de pointe en matière de détection et de prédiction de la maladie d’Alzheimer. De plus, notre approche basée sur les graphes améliore de 4 points de pourcentage en terme de précision de classification par rapport à nos meilleurs résultats obtenus par l’analyse d’hippocampe et présenté précédemment (voir table 1).

Par conséquent, dans cette deuxième partie, nous démontrons la nécessité de modéliser l’interdépendances des modifications en plus de capturer les modifications subtiles qui se produisent dans les structures clés et intégrer des fonctionnalités à différentes échelles anatomiques. De plus, nous avons également montré la complémentarité de notre nouvelle méthode basée sur l’imagerie avec les scores cognitifs souvent utilisés dans les essais cliniques pour diagnostiquer la maladie d’Alzheimer. Ces éléments mettent en évidence l’intérêt d’utiliser des méthodes basées sur l’imagerie en plus des scores cognitifs pour le dépistage précoce de la maladie d’Alzheimer.

Table 1: Comparaison des méthodes proposées dans cette thèse avec les méthodes de l'état de l'art utilisant des base de données similaires issues de ADNI1. Tout les résultats sont exprimés en pourcentage.

Méthode	Recalage	Caractéristique	CN vs. AD (ACC en %)	sMCI vs. pMCI (ACC en %)
(Coupé et al., 2012b)	Affine	Intensité	88.0	71.0
(Liu et al., 2012)	NG	GM	90.8	–
(Tong et al., 2014)	Affine	Graph	89.0	70.4
(Moradi et al., 2015)	NG	GM	–	74.7
(Tong et al., 2017a)	NG	Intensité	–	75.0
(Suk et al., 2017)	NG	GM	91.0	74.8
TBG (Hett et al., 2018d)	Affine	Texture	91.3	72.0
GHSG (Hett et al., 2018a)	Affine	Graph	–	74.7
GBSG (Hett et al., 2018b)	Affine	Graph	–	75.5
MGS (Hett et al., 2018c)	Affine	Graph	91.6	76.0

GM = Matière grise

NG = Non rigide

Contents

General introduction	1
I Computer-aided diagnosis for Alzheimer’s disease	11
1 Computer-aided diagnosis pipeline	14
1.1 Introduction	15
1.2 Data preprocessing	16
1.3 Segmentation of anatomical structures	19
1.4 Features extraction	20
1.5 Dimensionality reduction	20
1.6 Classification	21
1.7 Evaluation	22
1.8 Conclusion	23
2 MRI-based methods	25
2.1 Volume and surface-based methods	26
2.2 Voxel-based methods	29
2.3 Patch-based methods	30
2.4 Deep-learning methods	32
2.5 Conclusion	34
3 Patch-based grading	36
3.1 Patch-based methods	37
3.2 Patch extraction	40
3.3 Optimized patch-match	41
3.4 Influence of the parameters	42
3.5 Conclusion	47
Part conclusion	48
II Advanced hippocampus biomarkers	52
4 Adaptive fusion	56
4.1 Introduction	57

4.2	Background	57
4.3	Method	61
4.4	Conclusion	62
5	Texture-based grading	63
5.1	Introduction	64
5.2	Materials	65
5.3	Experiments	65
5.4	Results	68
5.5	Discussion	73
5.6	Conclusion	76
6	Multimodal-based grading	77
6.1	Introduction	78
6.2	Materials	79
6.3	Experiments	79
6.4	Results	82
6.5	Discussion	86
6.6	Conclusion	86
7	Hippocampal subfields grading	88
7.1	Introduction	89
7.2	Materials	90
7.3	Experiments	92
7.4	Results	93
7.5	Discussion	95
7.6	Conclusion	98
	Part conclusion	101
III	Modeling of Alzheimer’s disease signature	102
8	Graph of structures grading	106
8.1	Introduction	107
8.2	Background	107
8.3	Method	108
8.4	Conclusion	110
9	Graph of the hippocampal subfields	112
9.1	Introduction	113
9.2	Materials	113
9.3	Experiments	114
9.4	Results	115

9.5 Discussion	117
9.6 Conclusion	118
10 Graph of the brain structures	120
10.1 Introduction	121
10.2 Materials	121
10.3 Experiments	122
10.4 Results	123
10.5 Discussion	125
10.6 Conclusion	127
11 Multiple graph of brain structures	128
11.1 Introduction	129
11.2 Materials	129
11.3 Experiments	130
11.4 Results	132
11.5 Discussion	136
11.6 Conclusion	137
Part conclusions	138
General conclusion and perspectives	139
Appendix A: Publications	145
Appendix B: Classification methods	147
Bibliography	154
List of acronyms	179

List of Tables

1	Comparaison des méthodes de l'état de l'art	xiii
2.1	Comparison of state-of-the-art methods based on similar ADNI dataset	35
3.1	Description of the data used for the evaluation of the optimal parameters for the optimized patch-based grading. All data are from ADNI1 dataset.	44
5.1	Description of the dataset used in this work	65
5.2	Comparison of grading features based on histogram representation	69
5.3	Comparison of grading features based on mean of the grading values	70
5.4	Comparison with state-of-the-art methods	74
6.1	Description of the dataset used in this work	80
6.2	Mean AUC of the different features estimated over the whole hippocampal structure	83
6.3	Correlation of MPBG biomarker with cognitive scores	84
6.4	Comparison of our proposed MPBG biomarkers with state-of-the-arts methods	85
7.1	Correlation between MPBG biomarkers and cognitive scores	95
7.2	Comparison of our proposed MPBG biomarkers with state-of-the-arts methods	96
7.3	Comparison of multimodal patch-based grading and texture-based grading	97
9.1	Description of the dataset used in this work	114
9.2	Comparison of different hippocampal PBG approaches	117
9.3	Comparison with state-of-the-art methods based on the hippocampus	118
10.1	Description of the ADNI dataset used in this work.	121
10.2	Classification of sMCI versus pMCI	124
10.3	Comparison of the proposed method with state-of-the-art approaches	125
11.1	Description of the dataset used in this work	130
11.2	Comparisons of the different PBG approaches	134
11.3	Comparison with state-of-the-arts methods	135
11.4	Comparison of our graph-based approach with cognitive test scores	135
11.5	Comparison of state-of-the-art methods based on similar ADNI dataset	142

List of Figures

1	Pie-chart representing the prevalence of the main dementias	2
2	Illustration of the structural alterations caused by AD	3
3	Slices of MRI that represent brain structures	4
4	Schema of precession movement with magnetic field	8
5	Illustration of the pathological load decrease with the progression of AD	9
1.1	Schema of an usual computer aided diagnosis pipeline	15
1.2	Illustration of denoising method	17
1.3	Illustration of inhomogeneity correction	18
2.1	Illustration of the cortical thickness approach	28
2.2	Overview of the voxel-based morphometry methods	29
2.3	Illustration of the patch-based grading framework	31
2.4	Illustration of an usual CNN architecture	32
2.5	Illustration of an usual deep-learning approach based on patches	33
3.1	Illustration of the multiple instance learning graph	39
3.2	Illustration of the propagation step of the patchmatch procedure	41
3.3	Steps of the optimized patchmatch approach	43
3.4	Graph representing the evolution of classification performance with the increase of training library	44
3.5	Illustration of the evolution of grading maps throughout the increase of training library	45
3.6	Graph representing the evolution of classification performance with the increase of training library	46
3.7	Illustration of the evolution of grading maps throughout the increase of training library	47
3.8	Schema of the usual computer-aided diagnosis pipeline	48
4.1	Illustration of the general framework usually applied in multimodal image analysis method	58
4.2	Schema representing the early estimator fusion	60
4.3	Schema representing the pipeline of the proposed adaptive patch-based grading fusion scheme	61
5.1	Proposed adaptive fusion of texture-based grading framework	66
5.2	Sample of absolute value of complex number computed by Gabor filters	67

LIST OF FIGURES

5.3 Evolution of the accuracy and AUC according the number of direction 69
5.4 Comparison of different fusion schemes 71
5.5 Comparison of classification performances for different texture filters . 72

6.1 Proposed multimodal patch-based grading framework 82
6.2 Results obtained for different severities of cognitive impairments. From top to bottom slices on the coronal plane of the segmentation maps and the fusion of T1w and MD patch-based grading with the proposed multimodal patch-based grading method. The blue and the red colors represent the healthy and altered tissues, respectively. . . . 83

7.1 Segmentation of the hippocampal subfields 90
7.2 Proposed multimodal patch-based grading framework 91
7.3 Results obtained for different severities of cognitive impairments . . . 93
7.4 Distribution of the volume, MD and MPBG 99
7.5 Comparisons with the different considered biomarkers in each hippocampal area 100

8.1 Scheme of the proposed graph of structures grading pipeline 109

9.1 Illustration of the graph construction 115
9.2 Graphic representing the coefficients estimated by EN method 116

10.1 Illustration of the graph construction method 122
10.2 Chart of the most discriminant graph components 126

11.1 Schema of the proposed hierarchical graph of structure grading method 132
11.2 Schema of the multiple graphs of structure grading method 133
11.3 Illustration of the separating hyperplane computed with SVM 149
11.4 Illustration of the feature space partition conducted with the decision tree method 150
11.5 Illustration of a single hidden layer neural network 152

General introduction

Dementia

Dementias are characterized by a loss of thinking abilities that are severe enough to interfere with daily life. Symptoms of dementia can vary significantly by affecting memory, communication, language, ability to focus, reasoning, judgment, and visual perception. These symptoms are caused by damages of brain cells that interfere with the communications between the brains cells in different brain regions which are responsible for different cognitive functions. It has been estimated that 35.6 million people lived with dementia in 2010 and the number expected to almost double around 2030 (Prince et al., 2013). Dementia is not a part of normal aging and can occur at any age. However, the major part of people affected by dementia is around 65 years old. There are several sources of dementia (see Figure 1). However, the main causes of dementia come from AD, dementia with Lewy Bodies (DLB), Frontotemporal dementia (FD), Parkinson’s disease (PaD), and cerebrovascular disease (CDB).

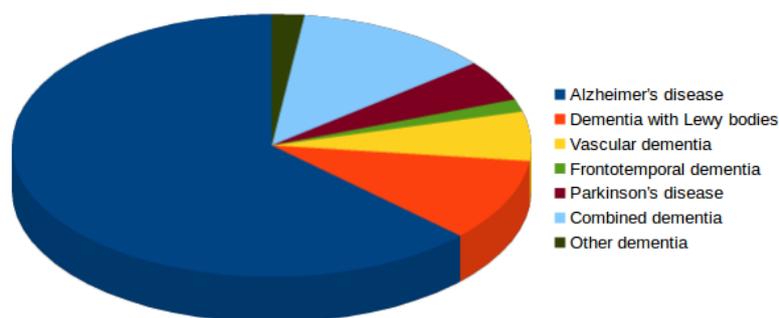


Figure 1: This pie-chart represents the proportion of the main dementias. As it is shown, patients suffering from Alzheimer’s disease represent around 62% of the dementias.

Alzheimer’s disease

Alzheimer’s disease (AD) is the most prevalent dementia affecting elderly people (Petrella et al., 2003). The prevalence of Alzheimer’s disease rises from 1% among people of 60 years of age to about 40% among people of 90 years old. With the constant aging of the population – according to the world health organization, the number of patients having AD will double in 20 years. Therefore, in the next decades AD will represent a substantial financial cost for the society (Wimo et al., 1997). AD leads to an irreversible neurodegenerative process causing mental dysfunctions such as long-term memory loss, language impairment, disorientation,

change in personality, and finally causes death (Alzheimer's, 2015). This disease is characterized by an accumulation of beta-amyloid plaques and neurofibrillary tangles composed of tau amyloid fibrils (Hardy, 2006) combined with synapse and neuronal loss. To date, no known therapy has been able to stop or slow down the progression of AD. Neuroimaging revealed that brain changes occur decades before the diagnosis is established. Moreover, the pathological load evolution is not linearly correlated to the cognitive function (see Figure 5). Thus, when the diagnosis of Alzheimer's disease is established, the pathological load is already high (DeCarli, 2003).

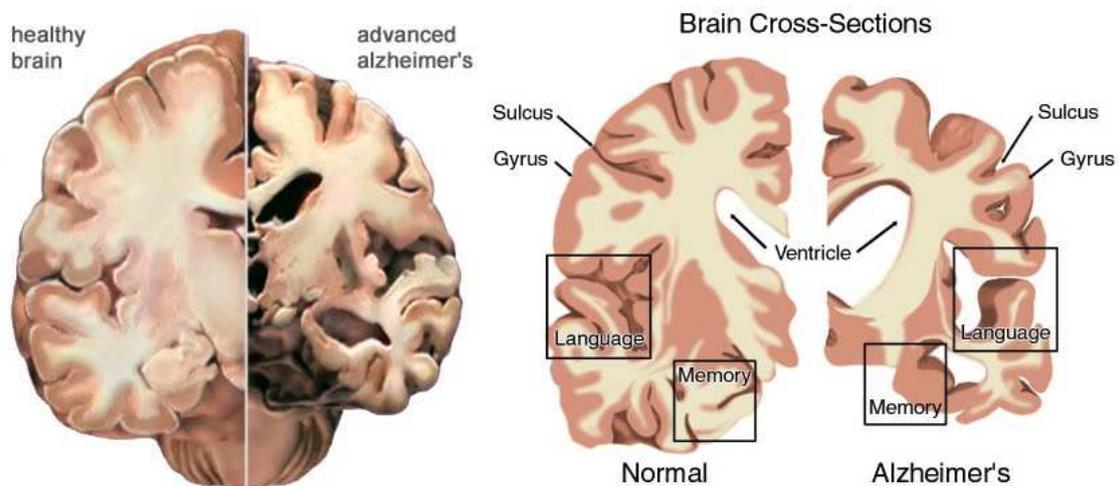


Figure 2: Illustration of the structural alterations caused by AD. On the left side, a brain from a cognitively normal subject and the other side a brain from a patient suffering from Alzheimer's disease. This illustrates the atrophy of global atrophy of the brain structure combined by hypertrophy of lateral ventricles and a focal atrophy of the medial temporal lobe that plays a central role into the conscious memory divided. Source: (Association et al., 2008)

Alzheimer's disease impacts brain structures at three different scales (see Figure 2). First, AD causes global atrophy of brain structures combined with a hypertrophy of the lateral ventricles. Second, AD causes diffuse atrophy of grey matter along the cerebral cortex. Finally, focal atrophy occurs. Pathological studies suggest that regions first affected in the medial temporal lobe by brain changes in typical disease progression are the entorhinal cortex (EC) and the hippocampus (Jack et al., 1992; Bobinski et al., 1999). Moreover, neuroimaging studies have shown that the hippocampus is the brain structure impacted by the most significant alterations at the early stage of AD (Frisoni et al., 2010; Schwarz et al., 2016).

Mild cognitive impairment

A prodromal phase of AD is mild cognitive impairment (MCI). The clinical symptoms of MCI are slight but the decreases of thinking abilities are measurable. MCI

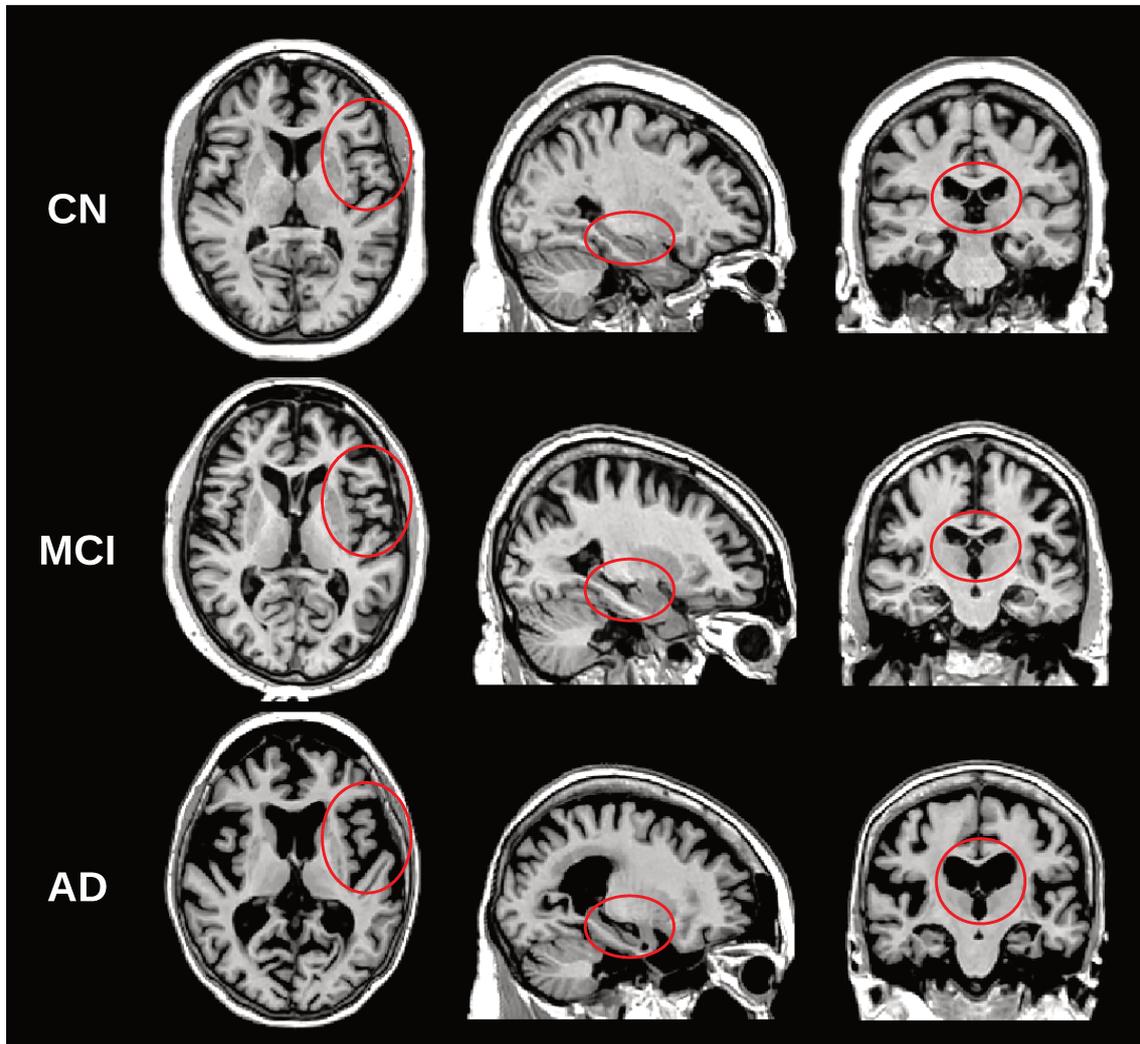


Figure 3: Slices of MRI that represent brain structures within, from left to right, axial plane, sagittal plane, and coronal plane. From the top the bottom, MRI from a cognitively normal subject, MRI from patients suffering from mild cognitive impairment and a patient who has Alzheimer’s disease. Although, the structural changes between cognitively normal subjects and patients suffering from Alzheimer’s disease are significant enough to be noticeable, MCI subjects present subtle brain alterations that are difficult to detect with a visual control. In red circle, brain structure having significant changes with the progression of AD; a part of the cerebral cortex (axial plane), the enthorinal cortex with the hippocampus (sagittal plane), lateral ventricles (coronal plane).

can be caused by heterogeneous sources of factor such as, in most of the cases, cerebrovascular disease, Lewy body dementia, frontotemporal dementia, primary progressive aphasia (Dubois and Albert, 2004). Moreover, previous studies have suggested that approximately 12% of subjects suffering from MCI progress to AD in the four years after the first MCI symptoms (Petersen et al., 1999). However,

although MCI subjects present a high risk of AD development, subjects suffering from MCI can remain stable (*i.e.*, do not convert to AD or any other dementia). This group of patient is named stable MCI (sMCI) on the opposite of patients who convert to AD in the following years after the diagnosis of MCI symptoms, named progressive MCI (pMCI). Finally, subject suffering from MCI can also convert to another dementia, or even recover to a cognitively normal status.

Diagnosis of Alzheimer's disease

The accuracy of the diagnosis is essential to ensure that patients receive appropriate treatments. Thus, to establish the diagnosis of Alzheimer's disease, clinicians conduct different tests. To assess the presence of Alzheimer's disease the physician evaluate: the impairment memory and cognitive skills, the exhibitions of changes in the patient personality or behaviors, the impact of cognitive impairment in daily life, and the potential cause of these symptoms.

Hence, the clinicians conduct different test such as neuropsychological tests that evaluate the cognitive and memory skills (Rogers et al., 1993), these scores are used to establish the degree of cognitive impairment. Among these tests we can cite (this list is not exhaustive):

Mini-mental state examination (MMSE) is a general test that was designed to evaluate the degree of cognitive impairment. It is used in clinical and research settings. MMSE consists of 30 points questionnaire that involves 9 categories: the orientation to time, the orientation to place, the registration, the attention, the recall, the language, the repetition, and the ability to execute complex commands (Folstein et al., 1975).

Alzheimer's disease assessment scale-cognitive subscale (ADAS-cog) was designed to measure the cognition, and it is frequently used in clinical trials. ADAS-cog measure the languages and memory, it involves 11 categories: the word recall task, naming objects and fingers, following commands, constructional praxis, ideal praxis, orientation, word recognition task, remembering test directions, spoken language, comprehension, word-finding difficulty (Mohs, 1983).

Rey auditory verbal learning test (RAVLT) was designed to assess immediate memory span, new learning, susceptibility to interference and recognition memory. The examiner reads a list of 15 words. Afterwards, the subject has to repeat all the words that he memorized. This process is carried out fives times (Schmidt et al., 1996).

Functional activity questionnaire (FAQ) is a neuropsychological test that assesses independence in daily life activities in normal aging or patient having mild dementia (Pfeffer et al., 1982).

Clinical dementia rating–sum of boxes (CDR-SB) is a numeric scale to measure the degree of severity of dementia symptoms. This score evaluates the cognitive performances in six categories: memory, judgment, problem solving, community affairs, home and hobbies, and personal care (Morris et al., 1988).

Although the neuropsychological tests are widely used in clinical routine, their abilities to provide an early diagnosis of Alzheimer’s disease are limited. Indeed, when the patients start to fail at these test, the neurodegenerative process caused by AD is already substantial. Early diagnosis of AD needs the development of new methods to detect the changes before these alterations decrease cognitive abilities of the patients.

Biomarkers for Alzheimer’s disease

A biomarker indicates the presence and an activity related to a specific disease. They are essential for the diagnosis and monitoring of the disease progression by quantifying the degree of changes in the structures or the function under analysis. Ideal biomarkers should be specific (*i.e.*, detect with a high accuracy subject that does not have the disease) and sensitive (*i.e.*, detect with a high accuracy patients suffering from the dementia).

Plasma biomarkers

Plasma is the liquid of blood where the red blood cells, white blood cells, and platelets are isolated. Plasma is easily extracted from blood by low-speed centrifugation combined with the presence of an anticoagulant. Several studies identified that the level of specific proteins differs from AD patients to cognitively normal subjects (Anoop et al., 2010). The main advantages of plasma biomarkers reside in their low-invasive aspect and the simplicity to obtain samples. However, limitations exist. First, plasma biomarkers are less correlated to AD. Moreover, their sensitivity and specificity for AD diagnosis are low compare to other biomarkers described below.

CSF biomarkers

Cerebrospinal fluid (CSF) is considered as a great source for biomarker development. CSF is in direct contact with the extracellular space of the brain and can reflect biochemical changes that occur inside the brain during the progression of AD. To date, three CSF biomarkers: beta amyloid ($A\beta_{42}$), total-tau (t-tau), and phosphorylated-tau (p-tau) have shown a high potential of diagnosis. CSF has several advantages, it can directly be correlated with AD, it is highly sensitive and specific. However, this

techniques is highly invasive. Indeed, samples of CSF are collected by lumbar puncture. Moreover, the diagnosis made with CSF is irreproducible due to the sample storage.

Imaging biomarkers

With the recent improvement of medical imaging techniques that provide powerful and non-invasive *in-vivo* imaging tools, new imaging-biomarkers were proposed. Those biomarkers can be grouped with respect to what kind of physical information they extract. Biomarkers based on magnetic resonance imaging (MRI) mainly capture structural alterations of brain structures such as hippocampus atrophy and lateral ventricles hypertrophy. Nuclear biomarkers using nuclear imaging such as Positron emission tomography (PET), capture the presence of specific protein such as abnormal beta-amyloid depositions in the hippocampus. Finally, functional biomarkers were developed to describe the alterations of cognitive function with functional imaging.

MRI-based biomarkers are *in-vivo* techniques with several advantages. They are non-invasive, they provide structural and functional details of the brain almost immediately, they are sensitive and specific. Furthermore, imaging biomarkers provide reproducible diagnosis. Therefore, even though this kind of biomarker is expensive and requires experienced personnel, the development of new imaging biomarkers has been highly intensive these last years.

Indeed, it has been shown that machine learning methods have the potential to assist in identifying patients with AD by learning discriminative patterns from neuroimaging data. To date, biomarkers based on MRI have reached good performances for AD diagnosis but are still limited for AD prognosis with the detection of MCI patient whom progress to dementia. Therefore, in this thesis, we focused on imaging biomarkers based on MRI with the aim to improve structural alteration detections.

Magnetic resonance imaging

Magnetic Resonance Imaging is a medical imaging technique that creates three-dimensional *in-vivo* images. On the opposite of computed tomography (CT) and PET scans, MRI is non-invasive and does not irradiate the body. MRI provides images with a high contrast of soft tissues (*i.e.*, brain, nerves, skins, fat, etc.). It is based on the principle of nuclear magnetic resonance (NMR). MRI requires a strong and stable magnetic field \vec{B}_0 produced by a superconducting magnet that aligns magnetic spin moments of protons. Weaker oscillating magnetic fields, called radio-frequency, are applied to slightly modify this alignment and to produce a precession phenomenon which provides a measurable electromagnetic signal. MRI consists of precisely locating the origin of this NMR signal by applying non-uniform

magnetic fields named gradients, which induce subtle different precession frequencies depending on the position of the protons. Nuclear magnetic resonance exploits the fact that the nuclei of certain atoms have a magnetic moment of spin. This is particularly the case of the hydrogen atom which is an element of water composition that is very abundant in biological tissues.

The spins of atomic nuclei are like spinning rotors around their axis z aligned with \vec{B}_0 (see Figure 4) making a fast movement of precession around the axis of the magnetic field. This movement is named Larmor's precession.

As the magnetic moments return to the direction of the static field z , the oscillating signal decreases until it disappears when all the magnetic moments are again aligned longitudinally (*i.e.*, in the z direction). The time that the nuclear magnetic moments have to regain their longitudinal alignment is called longitudinal relaxation time or T1 relaxation.

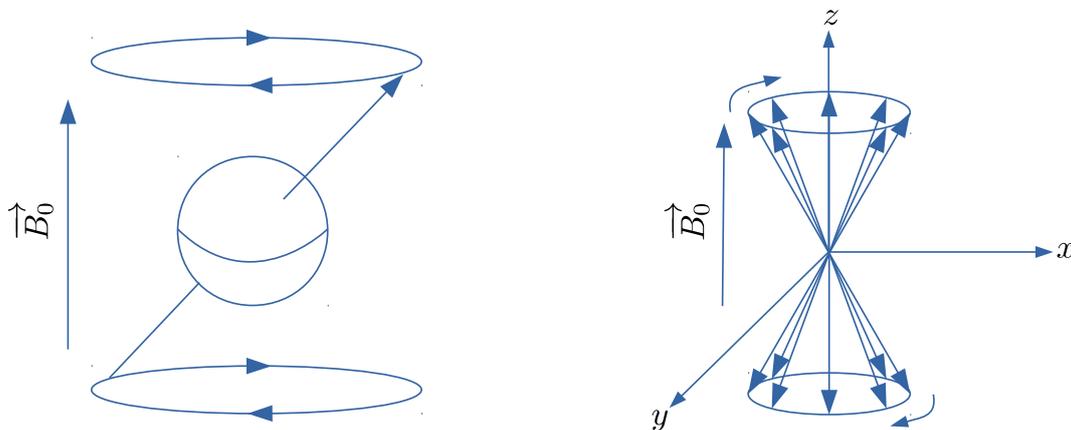


Figure 4: Schema of precession movement with magnetic field. At the left the schema representing the spin of a proton and at the right a group of spins. This schema is inspired from the illustration of the thesis of Vincent Noblet.

Moreover, the molecular agitation also contributes to another phenomenon. Whereas in theory, the magnetic moments should all rotate coherently around the axis z , (*i.e.*, with a different constant phase), the molecular agitation causes a heterogeneous physicochemical environment. Therefore their Larmor's frequency is also not perfectly equal to the theoretical Larmor's frequency. This results in a decrease in the signal related to their synchronous rotation over time. This phenomenon is named transverse relaxation time or T2 relaxation.

Structural MRI : By using a short repetition time and a short echo time (neutralizes T2 time differences), we obtain a T1-weighted image contrast, also called anatomical or structural MRI. In the T1-weighted images of the brain,

the white matter appears lighter than the gray matter. The cerebrospinal fluid located around the brain and in the lateral ventricles appears darker.

Diffusion MRI : Diffusion gradient techniques consist in measuring the Brownian motion of the water molecules in the tissues. This allows to extract information on the inhomogeneities of the tissues and in particular in the white matter. To achieve it, the diffusion measurements are performed on a higher number of directions (from 6 to more than one hundred) which make it possible to compute diffusion tensors in each voxel. From there, it is possible to define the main directions of the fibers that pass in each of the voxels and track them over the brain.

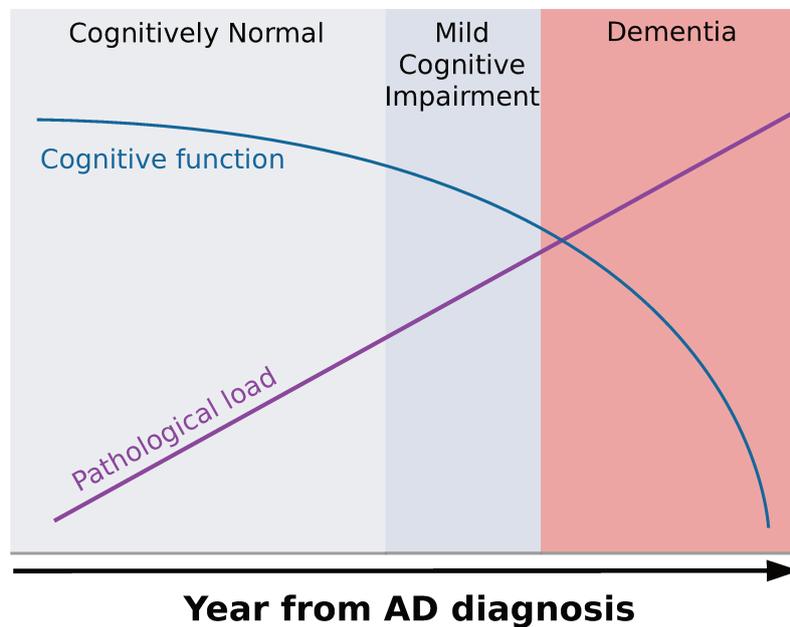


Figure 5: The pathological load linearly increase while the cognitive functions decrease exponentially. Therefore, the pathological load is high when the diagnosis is established. A prodromal phase of Alzheimer’s disease is mild cognitive impairment (MCI). However, a subject suffering from MCI can recovery to cognitively normal status, develop other dementias or convert to Alzheimer’s disease.

Challenges

As described in the previous section, numerous changes occur before the diagnosis of Alzheimer’s disease is established (see Figure 5). Among subjects suffering from MCI, the early stage of AD, subjects will remain in this state and never convert to AD, recover to cognitively normal status, or convert to others dementia. This

heterogeneity stresses the need to develop new methods to detect patients who will convert to dementia.

The early prediction of the subjects suffering from MCI symptoms who will convert to AD may improve the effectiveness of the future therapies by reducing the brain changes before the therapy starts. Moreover, the prediction of the conversion can also accelerate the development of new therapy by making the subject selection more accurate that would decrease the cost of clinical trials and would enable more accurate clinical studies.

The recent development of MRI techniques has enabled to study the anatomical changes such as brain structure atrophy with an high-contrast at a high-resolution. MRI data provides high-dimensional from the millions of voxels representing the brain tissues. To tackle the problem caused by the high-dimensionality of the problem, numerous methods have been proposed, some using a global context and others based on a selection of the most discriminative voxels. Although those methods have obtained good performances for the detection of AD, so far, the performances of the prediction of subjects who will convert to AD is not satisfying. The performance limitations of current methods are mostly due to the lack of abilities to capture subtle changes that occurs within the patients who will convert to AD, and a bad modeling of this changes through the brain. Thus, in this thesis, we propose new methods to increase the performances of conversion to AD prediction with biomarkers able to capture those subtle changes and with a novel modeling of the brain alterations.

Structure of this thesis

In this thesis, we have been interested in the development of new imaging biomarkers based on MRI. The first part of this thesis is a description of the elements composing a computer-aided diagnosis system and a state-of-the-art overview of MRI-based methods proposed in the literature. The second part focuses on hippocampus alterations detection with the improvement of patch-based grading technique. In this part, we propose a new adaptive fusion scheme that we applied to multi-textures and multi-modalities frameworks. In addition, we study the use of hippocampal subfields alterations with our new multimodal framework. In the third part, we describe a new graph-based method to better model the topology of alteration caused by AD usgin patch-based grading techniques. These methods have been applied with two anatomical representations, the hippocampal subfields, and the whole brain structures, demonstrating in both cases a great improvement of prediction performances. Moreover, we also proposed a new method to combine the graph produced by this two brain definitions into a unified framework that increases AD detection and prediction performances. Finally, this thesis ends with a general conclusion and perspectives of future works.

Part I:

Computer-aided diagnosis for Alzheimer's disease

In this part, we will describe the state-of-the-art computer-aided pipeline for the diagnosis of Alzheimer's disease. We will discuss of each step from the preprocessing to the classification methods that use features extracted from images to make a diagnosis decision.

Introduction

With the development of computer science and medical imaging techniques, the computer-aided diagnosis has become a helpful tool that helps clinicians for establishing diagnosis and prognosis by improving its accuracy. Indeed, the combination of magnetic resonance imaging and the last machine learning techniques has enabled to detect subtler anatomical changes at the earliest stages of dementia. This yields to an increase of the precision for Alzheimer’s disease diagnosis and prognosis.

In this part, we present the main elements that composed the most recent computer-aided diagnosis pipelines. Indeed, even though this thesis focuses on the feature extraction, many steps are mandatory to reach satisfying performances. To avoid bias in the extraction of feature, a proper preprocessing method must be designed. The preprocessing step enables a better intra-subject and inter-subject voxel value comparisons. Afterward, several choices have to be made, from the extraction of feature related to the information of anatomical changes to the classifier methods, numerous elements can impact the final classification results.

In this part, to help the reader of this thesis of understanding the motivation of the choices that we made in this thesis. We will describe the methods of each presented element.

Key-words

Computer-aided diagnosis, Classification pipeline, Alzheimer’s disease classification, Feature extraction, Patch-based grading method

Contents

- The Chapter 1 presents the general framework of computer-aided diagnosis. We describe in this chapter the different steps and the importance of each one in the final results. The different methods involved in common MRI preprocessing are presented. Also, the evaluation procedure and a description of the main datasets used to compare the methods proposed in the literature are provided. Finally, the different validation procedures are described.
- In the Chapter 2, an introduction of the state-of-the-art MRI-based methods is provided. Methods are grouped with the contextual information that they capture. ROI-based, voxel-based, deep-learning and patch-based methods are presented with their classification performances. Moreover, in addition, a table comparing all the mentioned methods is provided at the end of this chapter.
- The Chapter 3 presents three different patch-based grading methods. Indeed, in this chapter we present from the original to the last methods proposed in

the literature. The patch extraction step is also detailed with the introduction of a new optimized patch match method that improves the computational efficiency of the method. This chapter ends with a study of the parameters of the optimized patch match method and its influence for classification performances.

Chapter 1

Computer-aided diagnosis pipeline

1.1	Introduction	15
1.2	Data preprocessing	16
1.2.1	Denoising	16
1.2.2	Inhomogeneity correction	17
1.2.3	Registration	18
1.3	Segmentation of anatomical structures	19
1.4	Features extraction	20
1.5	Dimensionality reduction	20
1.6	Classification	21
1.7	Evaluation	22
1.7.1	Dataset	22
1.7.2	Validation	23
1.8	Conclusion	23

1.1 Introduction

The computer-aided diagnosis pipelines developed to detect AD are various and their architecture are related to the descriptor of the structural brain changes that methods capture. The pipelines are composed of different steps that usually involved data preprocessing, features extraction, eventual correction of feature bias, dimensionality reduction, and classification step; each step having their importance in the final results. The first step of the pipeline, named data preprocessing, aims to reduce the variabilities due to the acquisitions process and subject characteristics. This step is primordial to ensure the reliability of the results by avoiding any bias that could modify the evaluation of the results. Indeed, during the last decades, numerous methods were proposed for the detection and prediction of AD (Arbabshirani et al., 2017). The difficulty in comparing them comes from the large difference of preprocessing, dataset, and evaluation schemes (see Figure 1.1). In this chapter, we will discuss about the different steps composing the usual computer-aided diagnosis pipeline of MRI-based methods.

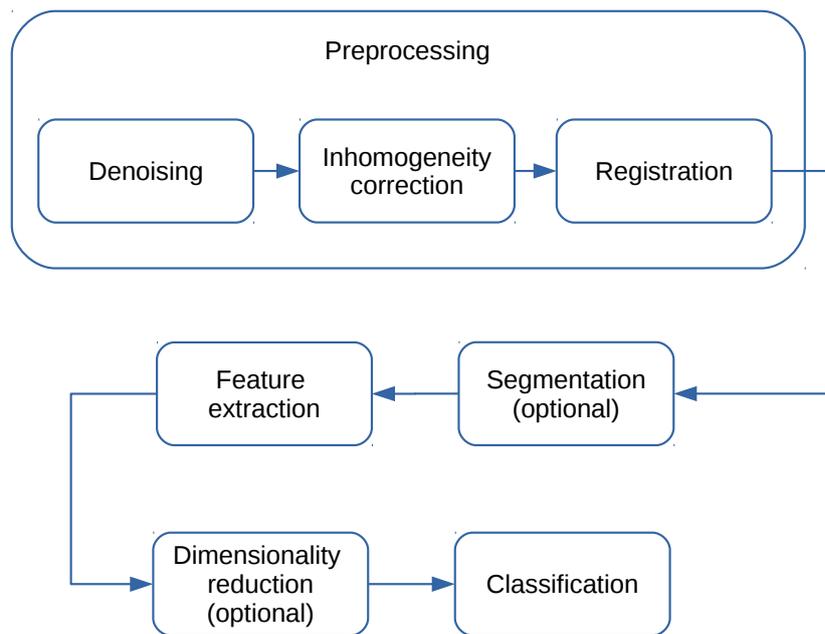


Figure 1.1: Schema of an usual computer aided diagnosis pipeline. The raw MRI given as input of the pipeline is preprocessed with a denoising, an inhomogeneities correction and a registration step. Afterwards, a segmentation of different brain structures are often performed. Then, feature extraction methods are applied. The features extraction can produces high dimensionality vector of features. Consequently some methods use dimensionality reduction method to fix this issue. Finally, the classification is performed.

However, several works have proposed to compare methods with the same pipeline

and the same dataset (Wolz et al., 2011; Cuingnet et al., 2011). The advantage of these works is that their results are easily interpretable. However, the main limitation comes from that only a few features are compared. Also, some studies have proposed to compare the impact of the different components of CAD pipelines with their predictive performances (Sabuncu et al., 2015; Samper-González et al., 2018). Recently, a standardized evaluation framework has been proposed for computer-aided diagnosis of dementia based on MRI data (Bron et al., 2015). This framework proposed to evaluate computer-aided diagnosis pipeline for three-classes classification task (*i.e.* CN, AD, and MCI). The results of this challenge have demonstrated the need for the use of similar dataset and robust validation procedures to compare the numerous methods proposed in the literature. Indeed, during the last decades a few open access MRI dataset related to Alzheimer’s disease have been released and it has become complicated to compare results of methods validating their approaches using different datasets.

In this chapter, to motivate the choices made in each contribution presented in this thesis, we introduce the different step that composes a usual computer-aided diagnosis pipeline, from the preprocessing step with denoising, inhomogeneity correction, registration and segmentation methods to the final evaluation. For each step, we present the advantage of limitation of each technique.

1.2 Data preprocessing

1.2.1 Denoising

MRI is corrupted by random noise that come from the acquisition process. The noise introduces an uncertainty in the measurement of voxels intensities that are used in further analysis. A straightforward technique to reduce the noise level would be to an average of multiple acquisitions in the MRI scanner. However, this technique is prone to numerous limitation and is not used in practice. The main limitation coming from the fact that averaging multiple acquisitions increase significantly the acquisition time. To avoid the augmentation of acquisition times numerous post-processing has been proposed in the literature (Geman and Geman, 1984; Saint-Marc et al., 1989; Perona and Malik, 1990; Black and Sapiro, 1999).

The most popular approaches are based on wavelet representation (Donoho and Johnstone, 1994; Donoho, 1995; Portilla and Simoncelli, 2003; Anand and Sahambi, 2010). Noise is characterized by high frequencies signal. Thus, the wavelet denoising approaches aim to delete wavelet coefficients where the noise energy is high. More recently, non-local mean filter (NLM) initially proposed for the denoising of natural images (Buades et al., 2005) have been adapted to process MRI. The approaches based on NLM assess that similar signals are present in several areas in the image. This group of methods aims to find a similar region, named patch, and average them to reduce the power of noise. On the opposite of the other methods, NLM methods

obtained excellent results for denoising while preserving the high frequencies in the images (Coupé et al., 2008; Manjón et al., 2008, 2010) (see Figure 1.2).

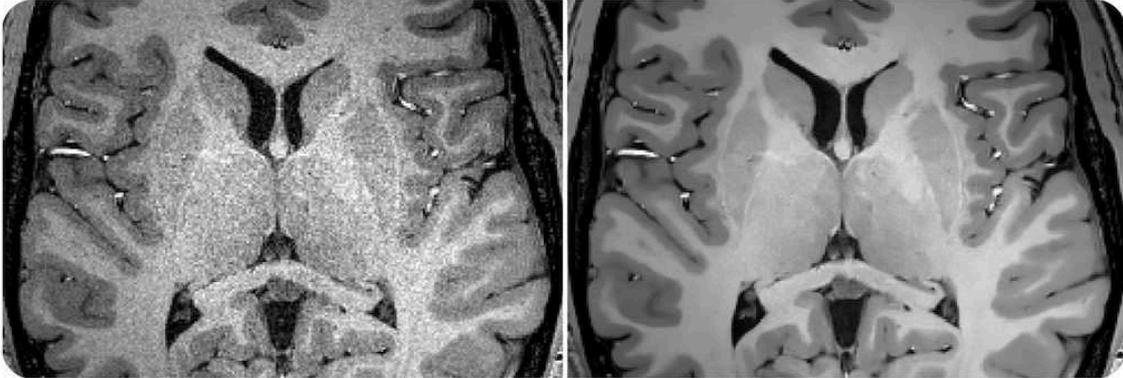


Figure 1.2: Illustration of denoising method, images at the left and right represent the data before and after the denoising process, respectively. Source: Coupé et al. (2008).

1.2.2 Inhomogeneity correction

One of the artifacts that MRI can suffer is the intensity inhomogeneity (see Figure 1.3). This artifact comes from the imperfection of the image acquisition. Intensity inhomogeneity can be seen as a smooth intensity variation across the image. This phenomenon causes a variation of the intensity of the same tissue according to its location. Moreover, the acquisition imperfection causing the intensity inhomogeneity can be grouped into two groups.

The first group is related to the properties of the MRI device. It includes static field inhomogeneity, bandwidth filtering of the data, eddy currents driven by field gradients, and radio frequency transmission and reception inhomogeneity. To correct this group of inhomogeneities, shimming techniques with particular imaging sequences, or by a MRI calibration using a phantom image are usually used (McVeigh et al., 1986; Axel et al., 1987; Wicks et al., 1993; Liang and Lauterbur, 2000).

The second group of inhomogeneities is related to the property of the patient himself. Indeed, his shape, position, and orientation inside the magnet and the specific permeability and dielectric properties of his body can modify the MRI intensity. To correct this group of inhomogeneity sources, numerous methods have been proposed (see (Vovk et al., 2007) for more details).

In this thesis, all MRI have been processed with an improved N3 bias correction to correct inhomogeneity (Tustison et al., 2010) that demonstrate state-of-the-art performances in many applications.

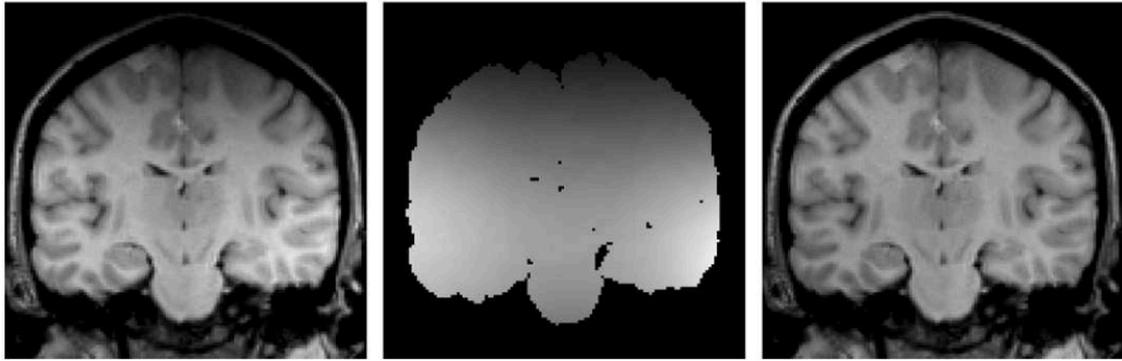


Figure 1.3: Illustration of inhomogeneity correction, from left to right: the original image without Inhomogeneity correction, the inhomogeneity field, and the inhomogeneity corrected image. Source: [Vovk et al. \(2007\)](#).

1.2.3 Registration

With the apparition of large datasets in the medical imaging field, the need of efficient method to wrap images into a same spatial space has risen. The aims of the registration method are to precisely aligned structures of interest coming from several subjects. Brain registration has been widely studied these last decades ([Collins et al., 1994](#); [Ashburner, 2007](#); [Liao and Chung, 2010](#); [Avants et al., 2011](#); [Mayer et al., 2011](#)). The registration methods can be grouped into two groups; the registration methods based on parametric models, such as nodes ([Wörz and Rohr, 2006](#)), curves ([Thompson et al., 1997](#)), or surfaces ([Davatzikos, 1997](#)) and the registration based on a voxel intensity. The advantage of parametric approach is the compact representation leading to a lower computational time compare to registration based on voxel intensity. However, the main limitation of parametric approaches comes from the imprecision related to the extraction of geometric primitive.

In addition, three different kind of transformation have been developed. The first is based on a rigid transformation. The rigid transformation consists of a linear mapping method that preserves distances, points, lines, and planes. The rigid transform is composed of a set of two operations, translation, and rotation. The second kind of transformation is composed of affine methods. Affine transform extends rigid transformation with shear and scale modifications. Moreover, as rigid methods, affine transform does not affect parallel lines which remain parallel after an affine transformation. The third kind of transformation is composed of non-rigid registration methods. On the opposite of rigid and affine transformations, non-rigid transformations are able to align images with a proper transformation at each voxel. However, although this last kind of transformations enable to better align anatomical structures, it is also less robust than rigid and affine transformations ([Avants et al., 2011](#); [Marstal et al., 2016](#)).

All along the experiments described in this thesis, we only use an affine registration method ([Avants et al., 2011](#)). Affine registration has two main advantage

compared to non-rigid method. First, the success rate is much higher and its computation time is lower than non-rigid registration that often need affine registration before processing estimation of non-rigid transformations. Indeed, affine registration uses only 12 parameters to deform the image with a translation, a rotation, and a possible scale modification and can be calculated in just a few seconds. Second, on the opposite of non-rigid transform, affine registration preserves proportion of structures volumes and does not change the shape and textures of anatomical brain structures.

1.3 Segmentation of anatomical structures

In computer-aided diagnosis pipeline, the segmentation step is often optional. However, the segmentation of brain structures on MRI images can be a mandatory step for many computer-aided diagnosis pipeline which are based on the analysis of specific ROIs (*i.e.*, cerebral cortex, entorhinal cortex, hippocampus, amygdala, etc.). To perform the extraction of ROIs, atlas-based methods are the most widely use strategy. Such methods use atlas to obtain definite representation of the brain structures with a strong a priori based on expert knowledge.

Among segmentation methods proposed these two last decades, three main atlas-based segmentation approaches have been successively developed. The first one is based on a non-linear registration of the atlas into the subject (Collins et al., 1995; Babalola et al., 2009). This approach has been widely used because of its robustness and its simple integration of expert knowledge *via* the deformation of manual segmentations. However, these methods can be prone to imprecision. Indeed, these methods can insufficiently capture the anatomical variability between subjects.

Significant improvement in segmentation performance has been achieved with the simultaneous use of several atlas (Rohlfing et al., 2004; Heckemann et al., 2006; Collins and Pruessner, 2010; Pipitone et al., 2014). Multi-atlas segmentation methods are based on an alignment of each atlas to the subject understudy with a non-rigid transformation. Once, each atlas are aligned within the subject understudy, a fusion method is applied to merge the information from each atlas (Iglesias and Sabuncu, 2015). In these methods, registration errors coming from inter-subject variability are considered as a random variable and are reduced with the use of several subjects.

Recently, the non-local means algorithm proposed by Buades et al. (2005) has been extended to patch-based label fusion (PBL) based on MRI of brain images (Coupé et al., 2011). This method has been proposed to reduce the calculation time of multiple atlas approaches based on non-linear registration. Rather than calculating non-linear transformations, the PBL method uses only an affine registration, and relies on the comparison of patches of voxels, or square neighborhoods, in a narrow search area between the subject to label, and several training atlas grouped into a library. This non-local strategy allows PBL to be robust to inaccuracies of

the registration step and effectively capture inter-subject variability. Thus, PBL improves multiple atlas methods based on global associations of labeled images. Moreover, patches locally describe the voxels content in the subject based on the information embedded in the atlases. This method produces state-of-the-art results for segmentation within reduced computing time (Giraud et al., 2016).

1.4 Features extraction

After the correction of the variabilities inherent to the acquisition process and subject characteristics, methods extract features describing the assessments related to the brain alterations caused by AD. The role of the feature is to provide quantitative information on the alterations caused by the disease. There are three categories of features depending on the context. The first category of the feature has been developed to analyze ROIs at a macroscopic level. Those features extract parameters related to the structures under analysis, for instance, volume, shape, thickness, etc (Wolz et al., 2011). The second category is composed of features working at a finer grained level. Those features analyze the modifications of tissues at a voxel scale (Ashburner and Friston, 2000). Finally, the last category is composed of intermediate features that embed the neighbor information named patch (Coupé et al., 2012b). Because the development of features is the main subject of this thesis, the following chapter will provide further description of the state-of-the-art methods of these three kinds of features (see Chapter 2).

1.5 Dimensionality reduction

Some feature extraction methods produce a considerable amount of features; it is especially the case for methods working at a voxel scale. The high dimensionality of the feature vectors does not enable to obtain satisfying performances. Therefore, dimensionality reduction techniques are often used. A first of group is composed of methods based on dimension projection techniques which are named principal component analysis (PCA) (Jolliffe, 2011). These methods are based on a projection of the vector into a more efficient feature space. Among proposed methods to perform PCA we can cite; the single value decomposition (SVD), which is a canonical method to compute PCA (De Lathauwer et al., 1994), and eigenvalues decomposition (EIG) of the covariance matrix (Acharya et al., 2012). These two methods are based on a selection of projected dimensions represented by the greater eigenvalues. Alternating least squares (ALS) is a ranking algorithm designed to better handle missing values (Kuroda et al., 2011). A major limitation of the methods using these techniques come from that they failed to help for understanding the disease progression since their results are not directly interpretable.

Besides dimensionality reduction, feature selection methods have been widely investigated. Indeed, it has been demonstrated that feature selection enable to in-

crease classification performances with high dimensional features (Chu et al., 2012). Moreover, on the opposite to PCA techniques, features selection methods provide interpretable results that can help to understand disease progression by capturing key elements. One of the most used approach is based on methods that only look at the intrinsic properties of the data. For instance, an usual method used to select discriminant features is based on p-values estimated with statistical test (Ashburner and Friston, 2000). Other proposed methods are based on dictionary learning techniques. The method usually used are: the least absolute shrinkage and selection operator (LASSO) (Tibshirani, 1996), its extension named elastic net (EN) (Zou and Hastie, 2005), that adds a L_2 regularization term which results in a grouping effect helping to select highly correlated features, and the sparse logistic regression (SLR) (Ryali et al., 2010). On the opposite of methods based on statistical test, methods based on dictionary learning embed the correlation of features and can avoid redundancy of the information captured by the features. Finally, methods embedded into classification algorithms are also used. These group of methods usually used the provided weight scores or the separability criterion used within the classification methods.

1.6 Classification

Machine learning and statistical classification is a set of learning approaches which aim to learn decision boundaries from a features space in order to discover hidden data labels. In other words, the goal of classification methods is to use data features, in our experiments the data represent MRI of brains, to identify which class it belongs to. For instance, in binary classification the data belong to two classes, and the classification methods aims to determine which class the sample belongs to according a set of training samples. In this thesis, we only consider four classes; the group that represents CN subjects, sMCI, pMCI and the group of patients whom suffer from AD.

The use of classification method depends of whether the feature space is linearly separable or not and how the distributions are modeled (*i.e.*, normal distribution, non-parametric, etc). For linearly separable features space linear classifier often perform satisfying results. The linear discriminant analysis (LDA) is frequently used to perform classification with the assessment of normal distribution (Fisher, 1936; Lachenbruch and Goldstein, 1979). To deal with non-parametric model of feature distribution, other methods are often applied such as support vector machine (SVM) (Cortes and Vapnik, 1995), decision tree (Quinlan, 1986), and its extension random forest (RF) (Breiman, 2001). Recently, with the development of deep-learning (LeCun et al., 2015), neural network has been introduced performing good performance. The strength of this techniques is its ability to handle linear and non-linear task by decomposition the classification problem into a network of linear classifiers organized in layers. However, due to its high number of connection, neural network needs high

number of training data. For the interested readers, these different methods are described in the annex **B** of this thesis.

1.7 Evaluation

1.7.1 Dataset

Besides in-house database used in some works, to accelerate the development of new imaging biomarkers for the early detection of Alzheimer’s disease, some program has proposed to collect and distribute imaging data. These programs have succeeded in building relatively large datasets that have contributed to the acceleration of development of new imaging biomarkers. The three most used, and open access database cited in the literature are:

Alzheimer’s Disease Neuroimaging Initiative (ADNI) is a North American campaign launched in 2003 with aims to provide MRI, positron emission tomography scans, clinical neurological measures, and other biomarkers ([Jack et al., 2008](#)). ADNI is a longitudinal multi-site observational study of elderly people. ADNI involves subjects from 50 years old to more than 95 years old. This database provides images from patients grouped in 3 different categories. Cognitively normal (CN) that is composed of subjects having normal cognitive abilities. MCI that is composed of two subgroups since its second phase (ADNI2), early MCI (eMCI) and late (lMCI), and AD patients. The patients have been followed on several years and their pathological status is available for each visit. Thus, it is possible to know which and when patients convert to AD. Thus, to date, ADNI provides large database composed of patients followed up for more than 4 years. ADNI is the most used dataset with numerous methods proposed in the literature using this database for their evaluation and comparison to former proposed methods ([Weiner et al., 2013, 2015](#)).

Australian Imaging Biomarkers and Lifestyle (AIBL) is an Australian study launched in 2006. AIBL is a longitudinal study of cognition on around 5 years ([Ellis et al., 2009](#)). The minimum age of subjects required in this study is 60 years old. Data have been collected from two centers. The data are grouped into 3 categories which are named CN for cognitively normal subjects, MCI for subjects suffer from MCI at the baseline and AD for data coming from patients with AD diagnostic at the baseline. Similar to ADNI, the longitudinal study also provides the time of conversion to AD of CN and MCI patients.

Open Access Series of Imaging Studies (OASIS) is a project compiling and distributing MRI datasets acquired from MRI studies of Washington University. The MRI represent subjects across the adult life span age from 18 to 96 years ([Marcus et al., 2007](#)). OASIS provides anatomical MRI data and clinical assessments of CN subjects and AD patients.

1.7.2 Validation

It has been known for long date that the use of same data to train a model and evaluate its statistical performances yields to obtains over-estimated results (Larson, 1931). To fix this issue, cross-validation (CV) was proposed. The general idea of CV comes from the separation of the same dataset into training and testing data. Since, numerous CV procedure have been introduced, for instance: repeated leave-n-out (LNOCV) (Shao, 1993), balanced incomplete CV (Shao, 1993; Wolz et al., 2011), Monte-Carlo CV (Picard and Cook, 1984), controlled 50% vs. 50% (Cuingnet et al., 2011), etc. In this section, we will describe only the three most common validation procedure used in the literature of AD classification.

Stratified k-fold CV (KFCV) procedure consists of randomly dividing the database into k different subset (Geisser, 1975; Liu et al., 2012). While one subsample is used as testing data, the $k - 1$ other subsamples are used to train the method. To capture the inner variability produce by the random selection, this procedure iterate N times, and the measures of final performances are obtained by averaging each intermediate performances measures obtained at each iteration.

Leave-one-out CV (LOOCV) procedure consist of train the method with $n - 1$ samples, where n is the size of the dataset (Stone, 1974; Allen, 1974; Geisser, 1975). With the large dataset, LOOCV has become computationally expensive. However, it has been shown that LOOCV provides almost unbiased estimator (Cawley and Talbot, 2004).

Using another database for the training step is a popular way to avoid any bias or over-fitting issues that can be introduce by the use of the same database for train and evaluate the proposed method. For instance some works proposed to use ADNI database to train their proposed methods and evaluate their performances with external databases (Coupé et al., 2015; Li et al., 2018).

1.8 Conclusion

It has been shown in this section the construction of computer-aided diagnosis pipeline is complex. To enable an unbiased comparison of MRI data a proper pre-processing involving denoising, inhomogeneity correction, and registration is mandatory. Then, according to the kind of extracted feature, a segmentation method can be applied. Finally, the evaluation of the pipeline is conduct with the use of the CV procedure. The complexity of the pipeline producing the final decision yields to a difficult method comparison. These difficulties are mainly led by the use of different datasets. Consequently, we decided to use in our works the ADNI database since it is a widely used database. We compare the proposed methods with only using

similar ADNI dataset. Moreover, we always provide a comparison with a baseline method based on well-established features to ensure non-optimistic results.

Chapter 2

MRI-based methods

2.1	Volume and surface-based methods	26
2.1.1	Volume-based feature	26
2.1.2	Shape-based feature	27
2.1.3	Cortical thickness	27
2.2	Voxel-based methods	29
2.3	Patch-based methods	30
2.4	Deep-learning methods	32
2.4.1	Volume-based feature	33
2.4.2	Patch-based feature	33
2.5	Conclusion	34

2.1 Volume and surface-based methods

Alzheimer’s disease causes different modification of brain structures:

- A global atrophy combined with a hypertrophy of lateral ventricles resulting from an increase of the quantity of CSF in the lateral ventricles.
- A diffuse atrophy of the gray matter along the cerebral cortex.
- A focal atrophy focused on the medial temporal lobe.

Although quantitative measures of the global brain atrophy with volume or surface analysis do not provide discriminant enough measures, numerous methods showed that the study of diffuse and focal atrophy of gray matter structures provides a good measure for the help of diagnosis. Among those methods two groups can be cited: the first one is composed of methods using the volume of specific structures while the second is composed of methods measuring the cortical thickness.

To perform volume and cortical-thickness analysis, the proposed methods need reliable segmentation methods for the labeling of brain structures of interest and for the delineation of the surface of the cerebral cortex. The most popular segmentation method used are the statistical parametric mapping toolbox (SPM) (Ashburner and Friston, 2005), FSL (Jenkinson et al., 2012), and FreeSurfer (Fischl, 2012). Recently, a web segmentation system based on multi-atlas label fusion has been developed (Manjón and Coupé, 2016).

2.1.1 Volume-based feature

Hippocampus volume has emerged as a gold standard imaging biomarker for quantify the severity of AD and help to establish the diagnosis of Alzheimer’s disease (Kesslak et al., 1991; Jack et al., 1992). However, the volume of EC has also been proposed as an imaging biomarker for the detection of AD (Bobinski et al., 1999; de Leon et al., 2001). Although some studies showed that the hippocampus volume is most discriminant in term of diagnostic separability (Frisoni et al., 2010; Coupé et al., 2012b; Schwarz et al., 2016) and it is considered as an enrichment biomarker for select patients for clinical trials (Jack Jr et al., 2011), the merit of hippocampus volume versus entorhinal cortex has been debated (Dickerson et al., 2001; Tapiola et al., 2008). Besides hippocampus and entorhinal cortex, amygdala structures have also been studied and showed a good diagnosis separability for the Alzheimer’s disease (Ledig et al., 2018). A limitation of volumetric biomarkers comes from its correlation to the total intracranial volume (TIV). Therefore, to reduce the variability due to the brain size, the volume of structures of interest has to be TIV corrected.

A large-scale study based on 834 subjects, Wolz et al. (2011), using a segmentation method based on an expectation maximization (EM) (Van Leemput et al., 1999), the experiments conducted in this study shows that hippocampus volume obtains an accuracy of 81% for the diagnosis of Alzheimer’s disease (*i.e.*, classification

of the cognitively normal subjects and Alzheimer’s disease patients), and 65% for the prediction of Alzheimer’s disease (*i.e.*, the classification of MCI subjects who remain stable and subjects suffering from MCI that converted to Alzheimer’s disease in the following years). It is interesting to note that this study showed that the volume of the hippocampus was among four methods on ten that predicted the conversion to Alzheimer’s disease more accurately than a random classifier. In their comparison with volumetric measurements, [Coupé et al. \(2012b\)](#) also evaluated the performance of volume and EC which obtained 70% of accuracy for the diagnosis of AD and 59% of accuracy for its prediction. This results demonstrated the interest of hippocampus for the early detection of AD when using volume.

2.1.2 Shape-based feature

Despite volume measurement has provided good results with its application to the hippocampus, it does not provide a full description of the changes that occurs with the progression of AD. Thus, in addition to volumetric biomarkers, shape descriptors have been introduced to capture finer modification of the hippocampus. Two groups of shape-based methods has been proposed. First, shape descriptors based on a representation using spherical harmonics (SPHARM) has been proposed to capture the model of the hippocampus and capture shape changes ([Brechtbühler et al., 1995](#); [Gerardin et al., 2009](#); [Cuingnet et al., 2011](#)). Another kind of approach is based on statistical shape models ([Davies et al., 2002](#)). These methods capture the variability of the brain structure over the population, because they produce high dimensionality feature vector, a dimensionality reduction such as PCA methods is often applied ([Shen et al., 2012](#)). An advantage of shape-based features is that on the opposite of the volume-based features, the shape descriptors are not correlated to the TIV.

It has been shown that shape-based method based on the hippocampus increases AD detection compare to the volume with a classification accuracy of 88% ([Gerardin et al., 2009](#)). However, a comparative study showed that does not prove the merit of shape descriptor of the hippocampus does not improve the results compare to its volume for the prediction of AD with 50% of balanced accuracy against 65.5% for the hippocampus volume ([Cuingnet et al., 2011](#)).

2.1.3 Cortical thickness

The cortical thickness is the most common cortical biomarker. Cortical thickness features aim to measure the diffuse atrophy of gray matter along the cerebral cortex by estimating its thickness (see [Figure 2.1](#)). The first way to compute cortical thickness is to extract the surface of the cerebral cortex. The thickness represents the minimal distance between two surface vertices of the meshes representing the inner and outer surfaces ([Fischl and Dale, 2000](#); [Miller et al., 2003](#)). Another way to compute cortical thickness biomarker uses a voxel-based approach ([Hutton et al.,](#)

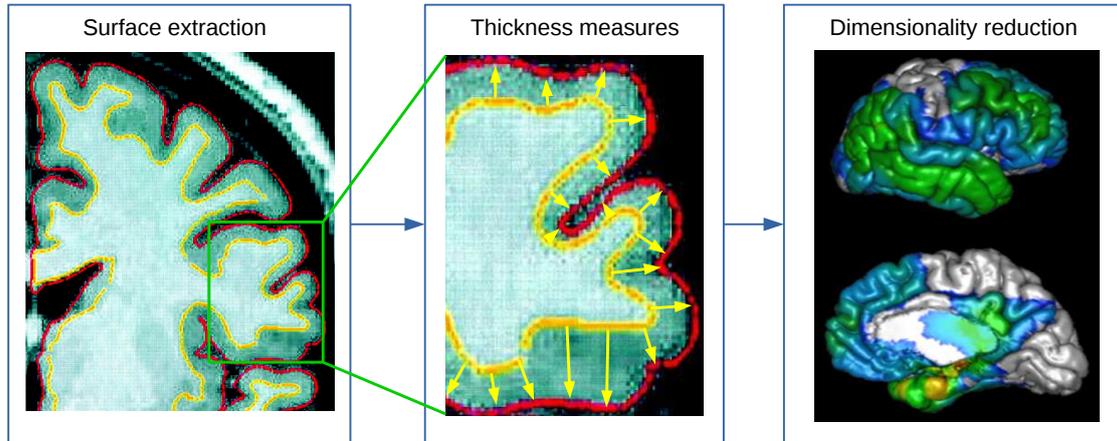


Figure 2.1: Illustration based on the images coming from the paper of (Fischl and Dale, 2000) and (Wolz et al., 2011). The yellow line represents the inner-surface and the red line represents outer-surface. From left to right, first the surface are extracted from the input image. The next step measure the normal distance between each vertex of the surfaces (yellow arrows). Finally, the cortical thickness are averaged into ROIs.

2008). In this method, the cortical thickness is defined at every voxel into the cerebral cortex as the length of the shortest path between the surface of white matter and CSF boundary. Cortical thickness methods produce highly dimensional features compared to the number of training dataset available. To reduce the dimensionality, selection or reduction methods are needed to avoid the over-fitting problem due to the curse of dimensionality. The most common method used to tackle this problem is based on a ROI approach, the mean cortical thickness is computed within each ROI and used as input of the classification step. An advantage of thickness measurements compared to volume is that cortical thickness is not correlated with the total intracranial volume (TIV) (Schwarz et al., 2016).

To date, several methods have been proposed to use cortical biomarkers for AD classification. An extensive method comparison (Wolz et al., 2011) estimated cortical thickness with a surface-based approach and averaged into defined ROIs. This approach obtained 81% of accuracy for AD detection and 56% of accuracy for AD prediction. A cortical thickness approach, using another surface-based extraction method (Eskildsen and Østergaard, 2006), obtains 85% of accuracy for AD detection and 63.7% of accuracy for AD prediction (Eskildsen et al., 2013). Querbes et al. (2009) proposed to adopt a decision tree classifier to better deal with the high dimensionality of cortical thickness feature, this method obtained 85% of accuracy for AD detection and 73% of accuracy for AD prediction. An approach proposed to capture inter-regional covariation with a thickness network integrated into a multi-kernel learning approach obtained 89% of accuracy for AD detection and 64% of accuracy for Alzheimer’s disease prediction (Raamana et al., 2015). Another method based on the same idea of constructing the thickness based network and focusing on the

prediction of AD obtained 75% of accuracy (Wee et al., 2013). These two methods demonstrated the interest of modeling the alterations relationship that occurs through the same subject.

2.2 Voxel-based methods

Voxel-based morphometry (VBM) methods (Ashburner and Friston, 2000) have been widely used for the identification of discriminant area into the brain at a voxel scale. The voxel-based method consists of a transformation of MRI data into the same space often using non-rigid transformation that correct individual brain shape differences. Afterward, a segmentation of gray matter, white matter, and CSF is conducted based on voxel intensities. To increase the reliability of the statistical analysis a smoothing of the images is done before the final comparison between group conducted at each voxel and based on the feature of interest (*i.e.*, gray matter or white matter). The result of this techniques is a statistical map that estimates at each voxel the statistical significance between considered groups (see Figure 2.2). In opposition to volume and surface-based methods, VBM approaches are not restricted to specific regions of interest. Moreover, VBM method is automated and usually independent of segmentation methods.

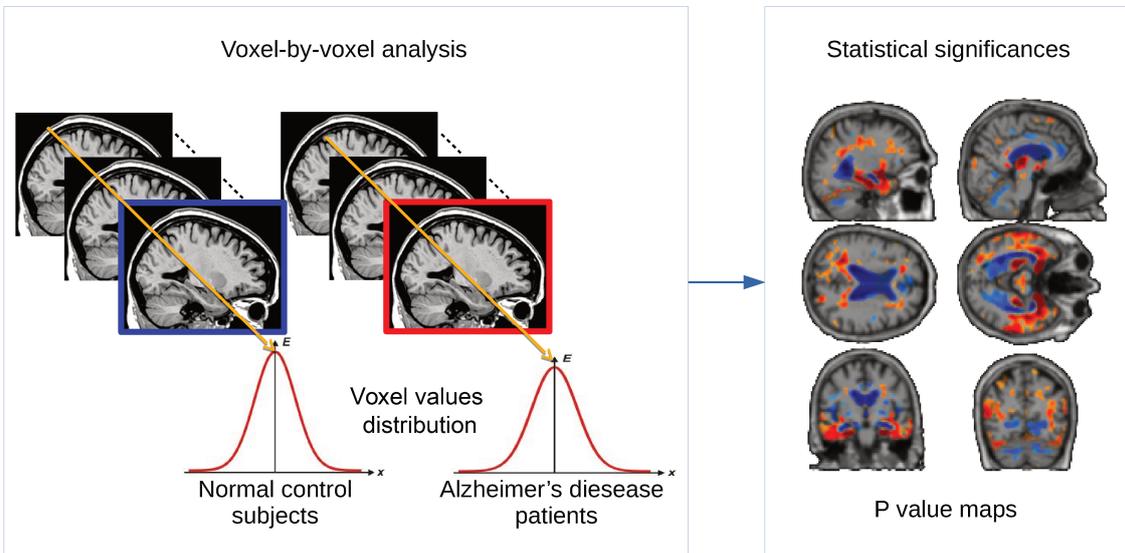


Figure 2.2: Overview of the voxel-based morphometry methods. After the images are registered into the same space with a non-rigid transformation, a statistical analysis is conducted based the distribution at each location to determine which voxels are significantly discriminant. This results a map of statistical coefficients.

Besides the use of the voxel values, some approaches proposed to use the deformation fields estimated at the non-linear registration step. Tensor-based morphometry

(TBM) (Ashburner et al., 2000; Hua et al., 2008) methods share the same principle than VBM. Instead of considering intensity or gray matter probabilities values, the TBM methods consider the Jacobian matrix at each voxel from the deformation field. In this kind of approach, the determinant of the Jacobian matrix is used to estimate the local atrophy. As VBM, the statistical maps are usually computed from the determinant of the Jacobian matrix and used to select the most discriminant voxels.

VBM and TBM studies have shown global cortical atrophy with focal atrophy of the sensorimotor cortex, occipital lobes, and cerebellum (Karas et al., 2003). Besides, their results have also shown that medial temporal lobe with the amygdala, entorhinal cortex and hippocampus suffer from the most significant loss (Baron et al., 2001; Hirata et al., 2005). These results confirm the interest of using ROI approaches focusing on structures belonging to the medial temporal lobes.

Wolz et al. (2011) used TBM approach to detect and predict AD. In their work, they obtained a detection accuracy of 87% and a prediction accuracy of 64%. Another voxel-based approaches used a SVM classifier (Abdulkadir et al., 2011; Cuingnet et al., 2013). These two methods have obtained similar accuracy with respectively 87% and 91% for the detection of AD.

More recently, a VBM approach aiming to improve the classifier method for the prediction of AD obtained 74.7% of accuracy Moradi et al. (2015). On the opposite of most of the previous VBM methods, the feature selection step has not been conducted with a statistical test. Indeed, this VBM approach use a sparse logistic regression method to capture the most discriminant voxels combined with a low density separation (LDS) method to obtain the final prediction.

2.3 Patch-based methods

The use of patch-based methods for Alzheimer’s disease detection was initially proposed by Coupé et al. (2012b) with the patch-based grading (PBG) framework. The main idea of this exemplar-based method is to use the capability of patch-based techniques to capture subtle signal alterations. The idea is to propagate local pattern information embedded in a template library to the image under analysis. The original method has been proposed to capture subtle alterations related to anatomical degradations caused by AD. To determine the pathological status of the subject under study, PBG method estimates at each voxel the severity of structural alterations by a non-local similarity measure.

In the next chapter, we will describe three kinds of patch-based methods, the first one based on the original PBG method named scoring of non-local image patch estimator (SNIPE) (Coupé et al., 2012b), the second one proposes an approach estimating a global patch-based grading feature based on a sparse representation (Tong et al., 2017a), and finally a graph-based approach that consists of capturing intra-subject variability information with a multiple learning approach (Tong et al.,

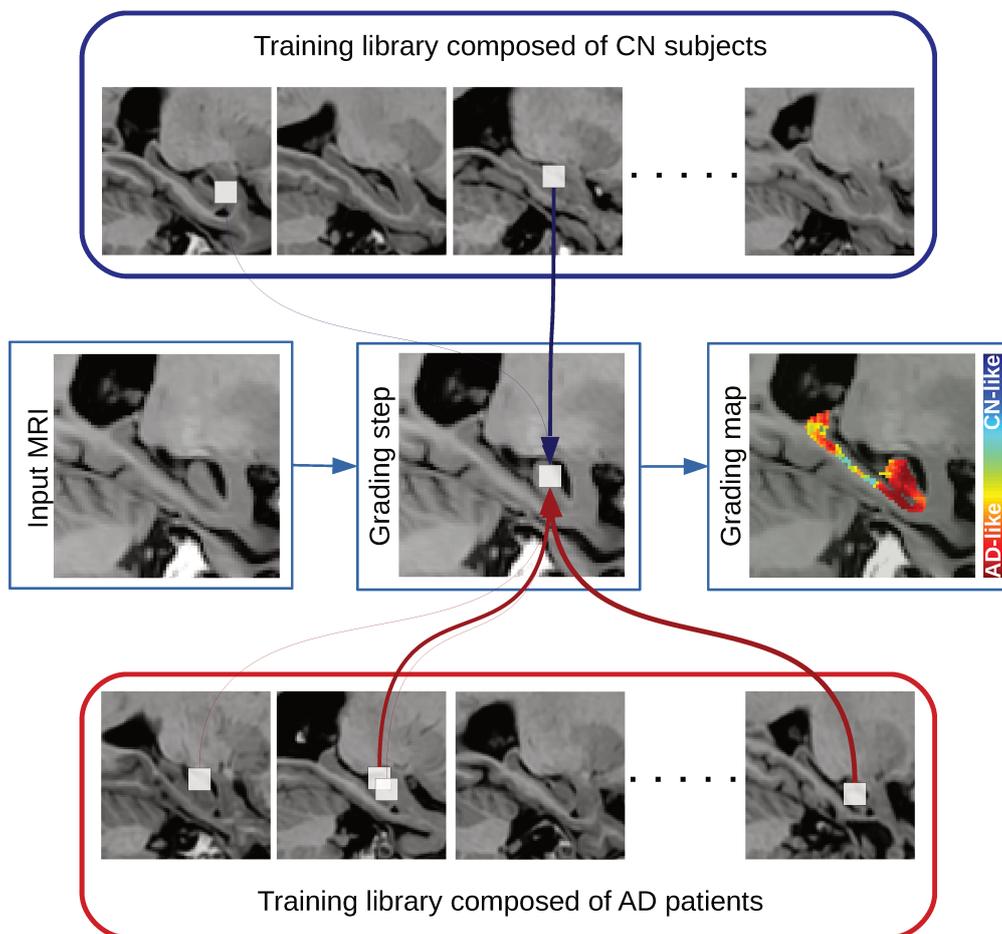


Figure 2.3: The main idea of patch-based grading framework is to capture local pattern similarities that are estimated on patches. Patch-based grading method provides at each voxel a value representing the probability of alterations. These probabilities are estimated by extracting the most similar patches in images from two templates library. One composed of cognitively normal subjects and the other one composed of patients suffering from Alzheimer’s disease. To estimate the grading value, a label related to the pathological status for each patch is set: to 1 for cognitively normal subject and -1 for Alzheimer’s disease patients. Here the input MRI comes from a AD patient.

2014).

SNIFE method increases the classification performances of methods based on an analysis of the hippocampus structure. Indeed, SNIFE obtained an accuracy of 88% for AD detection and accuracy of 71% for its prediction (Coupé et al., 2012b). This method has outperformed the previous approach based on the analysis of the hippocampus. The application of the sparse-based grading method to the hippocampus mask obtained an accuracy of 69% for AD prediction. However, the framework applied on the whole brain has obtained high performances with a prediction accuracy

of 75% (Tong et al., 2017a). Finally, the patch-based analysis based on a multiple instance learning approach has been evaluated on the detection and prediction of Alzheimer’s disease. For the prediction, Tong et al. (2014) obtained an accuracy of 89% and 70% for the prediction. These results are similar to those obtained with the original PBG approach (Coupé et al., 2012b).

2.4 Deep-learning methods

Deep-learning with its most popular form, the convolutional neural networks (CNN), has outperformed the state-of-the-art results in many visual recognition tasks. CNN architecture is usually based on a stack of several hidden layers (see Figure 2.4). Each hidden layers can be a convolutional layer, a pooling layer, a rectifier layer, or a fully-connected layer. Convolutional layers and pooling layers are designed to detect particular pattern and reduce the problem dimensionality. Rectifier layer is working as an activation function. Neurons in a fully connected layer have full connections to all activations in the previous layer, these layers are designed to conduct a high-level abstraction modeling.

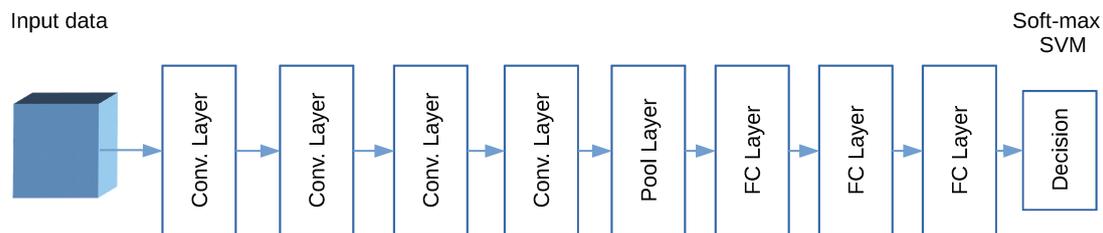


Figure 2.4: Illustration of an usual CNN architecture. From left to right: Several convolutional layers are applied to the input data, each convolutional layer apply different filters to detect particular patterns. Afterward, pooling layers are used to reduce the dimensionality. Finally, to learn model, fully-connected (FC) are trained. To take a decision, linear classifier such as soft-max or SVM are often used.

Thus, deep-learning can provide end-to-end methods that can use raw images and detect the most discriminant regions automatically via activation maps. However, so far this kind of framework has barely obtained competitive performances for Alzheimer’s disease classification (see CADDementia challenges for instance (Bron et al., 2015)). The main issue may be from the low number of training data available compared to the high number of parameters to optimize.

In this section, we describe two methods that proposed different approaches to deal with the issue coming from the limited size of the dataset. The first one uses gray matter volumes in ROIs (Suk et al., 2017) and the second uses bag of patches for data augmentation (Luo et al., 2017; Liu et al., 2018; Li et al., 2018).

2.4.1 Volume-based feature

The framework proposed by [Suk et al. \(2017\)](#) uses gray matter volumes as inputs of a CNN architecture and has significantly improved AD prediction. The strength of this method is to reduce the input dimensionality by providing relevant features to the CNN. To extract a high-level model of the alteration relationships between several brain structures, this deep volume-based method uses an iterate sparse regression model set up with different parameters. The sparse regression models produce a set of coefficients representing the different structures having the most discriminant structural modifications. This results in a matrix of discriminant features that are used as input of a CNN composed of two convolutional layers and two fully connected layers. This method has been validated on a large dataset and obtained 91.0% for the detection of AD and 74.8% for its prediction.

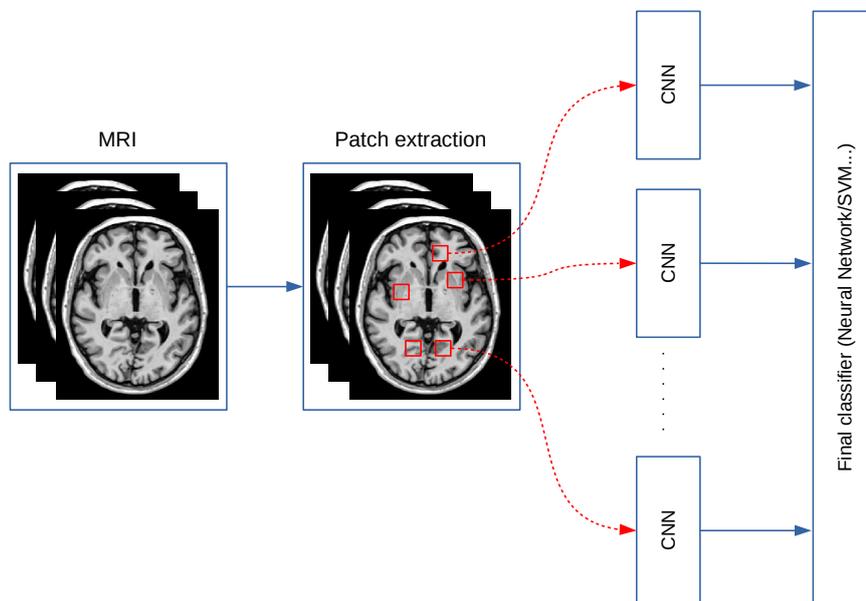


Figure 2.5: Illustration of an usual deep-learning approach based on patches. First, discriminant patches are extracted from the images. Then a CNN is trained for each patch. Finally, a final classifier use the output of each CNN to compute the final decision.

2.4.2 Patch-based feature

Most of the methods based on a straightforward application of well-known CNN architecture does not obtain satisfying results compared to state-of-the-art methods. However, some of them using a data augmentation procedure based on patches reach competitive performances for AD detection ([Luo et al., 2017](#); [Liu et al., 2018](#); [Li et al., 2018](#)) (see Figure 2.5). Among them, ([Liu et al., 2018](#)) proposed a landmark-based deep multi-instance learning approach. In opposite to end-to-end deep learn-

ing methods, this approach uses patches extracted from the most discriminant landmark with a regression forest approach for data augmentation. For each training data, numerous patches are extracted at each landmark and used as input of a CNN method. Each landmark represents a bag of patches, a CNN network composed of 6 convolutional layers and two fully connected layers are used for estimating the degree of alterations for each bag of patches. Then, the output of the last fully connected layers are concatenated and connected to 3 last fully connected layers and a soft-max function is used to take the decision (*i.e.*, AD or CN, and pMCI or sMCI). This landmarks-based neural network obtains state-of-the-art results for AD detection. However, even though the results in terms of accuracy are relatively high compared to the state-of-the-art methods with 76.9%, the testing dataset is poorly balanced, and the results in terms of balanced accuracy is lower with only 62.3%.

2.5 Conclusion

In this chapter, we presented the state-of-the-art MRI-based methods for the classification of Alzheimer’s disease. As described in table 2.1, these methods are grouped into 4 categories:

- Methods based on the analysis of volume and surface of specific brain structure such as the hippocampus and the enthorinal cortex.
- Voxel-based analysis methods that capture structural changes at a fined grained by comparing the value at a voxel scale.
- Methods based on deep-learning approaches that are based on high-level learning.
- Patch-based framework that is an exemplar-based approach which captures neighbor context information.

The patch-based framework provides a simple and efficient approach to extract relevant features from images. Indeed, the advantages of the patch-based framework over the other methods comes from its efficient way to capture subtle pattern changes using a local similarity measurement. On the opposite of volume and surface-based approaches, patch-based methods do not limit its analysis to a global structure representation. Moreover, it captures local information embedded in the patch and enable an inter-patches comparisons when voxel-based methods only enable to compare distribution values within the same voxel. Finally, a last benefit compare to deep-learning approaches is that patch-based framework need smaller dataset to perform state-of-the-art results.

2. MRI-based methods

Table 2.1: Comparison of state-of-the-art methods based on similar ADNI dataset. The methods are grouped into: volume and surface-based, voxel-based morphometry, patch-based and deep-learning approaches. The results are expressed in term of detection (*i.e.* cognitively normal subject vs. Alzheimer’s disease patients) and prediction (*i.e.*, non-converted MCI vs. converted MCI) of Alzheimer’s disease. Results are expressed in percentage of accuracy and balanced accuracy for results obtained with unbalanced dataset.

Method	Region	Feature	Classifier	Detection	Prediction
Volume and surface-based					
(Colliot et al., 2008)	Hipp	Volume	NN	84.0	-
(Querbes et al., 2009)	CC	Thickness	DT	85.0	73.0
(Gerardin et al., 2009)	Hipp	Shape	SVM	88.0	-
(Wolz et al., 2011)	Hipp	Volume	LDA	81.0	65.0
(Wolz et al., 2011)	CC	Thickness	LDA	81.0	56.0
(Cuingnet et al., 2011)	Hipp	Shape	SVM	76.5*	50.0*
(Coupé et al., 2012b)	EC	Volume	LDA	70.0	59.0
(Eskildsen et al., 2013)	CC	Thickness	LDA	85.5	63.7
(Wee et al., 2013)	CC	Thickness	SVM	-	75.0 [†]
(Raamana et al., 2015)	CC	Thickness	MKL	89.0	64.0
Voxel-based morphometry					
(Wolz et al., 2011)	Brain	Tensor	LDA	87.0	64.0
(Abdulkadir et al., 2011)	Brain	Gray matter	SVM	87.0	-
(Cuingnet et al., 2013)	Brain	Gray matter	SVM	91.0	-
(Moradi et al., 2015)	Brain	Gray matter	LDS	-	74.7
Patch-based					
(Coupé et al., 2012b)	Hipp	Patch	LDA	88.0	71.0
(Tong et al., 2014)	MTL	Patch	SVM	89.0	70.0
(Tong et al., 2017a)	Hipp	Patch	SVM	-	69.0
(Tong et al., 2017a)	Brain	Patch	SVM	-	75.0
Deep learning					
(Suk et al., 2017)	Brain	Volume	CNN	91.0	74.8
(Luo et al., 2017)	Brain	Patch	CNN	83.0	-
(Liu et al., 2018)	Brain	Patch	CNN	90.5*	62.3*
(Li et al., 2018)	Brain	Patch	CNN	89.5	-

* Balanced accuracy

[†] Reduce ADNI1 dataset

LDA = Linear discriminant analysis
LDS = Low density separation
SVM = Support vector machine
DT = Decision tree
CNN = Convolutional neural network
MKL = Multi kernel learning

MTL = Medial temporal lobe
EC = Entorhinal cortex
Hipp = Hippocampus
CC = Cerebral cortex
NN = Nearest neighbor

Chapter 3

Patch-based grading

3.1	Patch-based methods	37
3.1.1	Scoring of non-local image patch estimator	37
3.1.2	Sparse-based grading	38
3.1.3	Multiple instance learning	39
3.2	Patch extraction	40
3.3	Optimized patch-match	41
3.4	Influence of the parameters	42
3.4.1	Number of patches	44
3.4.2	Size of the training library	45
3.5	Conclusion	47

3.1 Patch-based methods

3.1.1 Scoring of non-local image patch estimator

In his original paper, [Coupé et al. \(2012a,b\)](#) proposed to capture of hippocampus structural alterations with a new scale of analysis. The scoring of non-local image patch estimator (SNIPE) extract similar patches and encode the distances between patches from training library and patch from the voxel under analysis into a weight function.

The method starts as follows, a training library T composed of two datasets of images is built: one with images from CN subjects and the other one from AD patients. Next, for each voxel x_i of the region of interest in the considered subject x , the PBG method produces a weak classifier denoted g_{x_i} . This weak classifier provides a surrogate of the pathological grading at the considered position. The weak classifier is computed using a measurement of the similarity between the patch P_{x_i} surrounding the voxel x_i belonging to the image under study and a set $K_{x_i} = \{P_{t_j}\}$ of the closest patches P_{t_j} , surrounding the voxel t_j , extracted from the template $t \in T$ (see [Figure 2.3](#)). The grading value g_{x_i} at x_i is defined as:

$$g_{x_i} = \frac{\sum_{t_j \in K_{x_i}} w(P_{x_i}, P_{t_j}) p_t}{\sum_{t_j \in K_{x_i}} w(P_{x_i}, P_{t_j})} \quad (3.1)$$

where $w(x_i, t_j)$ is the weight assigned to the pathological status p_t of the training image t . We estimate w such as:

$$w(P_{x_i}, P_{t_j}) = \exp\left(-\frac{\|P_{x_i} - P_{t_j}\|_2^2}{h^2}\right) \quad (3.2)$$

where $h = \min\|P_{x_i} - P_{t_j}\|_2^2 + \epsilon$ and $\epsilon \rightarrow 0$. The pathological status p_t is set to -1 for patches extracted from AD patient and to 1 for patches extracted from CN subject. Therefore, the PBG method provides at each voxel a score representing an estimation of the alterations caused by AD. Consequently, cerebral tissues strongly altered by AD have grading values close to -1 contrary to healthy one with scores close to 1 .

Finally, to deal with the variability of hippocampal volume and to reduce the feature dimensionality, the mean grading value is computed as follows:

$$\bar{g} = \frac{1}{|V|} \sum_{x_i \in V} g_{x_i}, \quad (3.3)$$

where \bar{g} is the mean grading value, V is the region of interest and $|V|$ is the number of voxels in this region. Thus, for each subject, PBG applied on the hippocampus provides one to two features (*i.e.*, if V represent the combined label of both hippocampus, or one hippocampus) considered as input of the classification method. It is interesting to note that the use of CN and AD data to estimate the structural

alterations similarity provides the best results to model the relationship between MCI and images from CN and AD. Then, the use of CN and AD to estimate grading values on MCI subjects provides better results than a straightforward approach using MCI to classify MCI subjects. Indeed, the results of PBG decrease when the grading values are computed with the group sMCI and pMCI. This effect could come from that MCI group is too noisy compared to hippocampus from CN and AD patients. Consequently, the uncertainty introduced with the use of MCI reduces the classification performances.

3.1.2 Sparse-based grading

To model the relationship between MCI and images from CN and AD patients, [Tong et al. \(2017a\)](#) proposed a dictionary learning approach combined with a sparse representation. Each MCI subject is assumed to lie in the space of training population, which means that it can be represented by a linear combination of CN and AD patients in the training population.

In opposite to the previous method that analyzes local alteration in a cubic fixed size patch, sparse-grading method proposes to study the similarity with a non-definite shape of patch. Indeed, the first step of this method is based on a feature selection proposed to capture the most discriminant voxel with a elastic-net regression ([Janoušová et al., 2012](#)). Thus, after features selection the N most discriminant voxels are determined and N intensity values are extracted from each image. This results in $X^{ADCN} \in R^{N \times M}$ the matrix containing the N intensities of M training images and $X^{MCI} \in R^{N \times 1}$ that contains the N intensity values of the test MCI image. The sparse-grading map is obtained by minimizing the following function:

$$\hat{\alpha} = \min_{\alpha} \frac{1}{2} \|X^{MCI} - X^{ADCN} \alpha\|_2^2 + \lambda_1 \|\alpha\|_1 + \lambda_2 \|\alpha\|_2^2 \quad (3.4)$$

where α represents the coding coefficients, λ_1 and λ_2 represent the hyper-parameters used to weight the L_1 and the L_2 norms, respectively.

$$g = \frac{\sum_{j=1}^N \hat{\alpha}(j) p_j}{\sum_{j=1}^N \hat{\alpha}(j)} \quad (3.5)$$

As described in the previous method, p_j represents the pathological status of the training image X_j^{ADCN} . p_j is set to -1 to images of patient suffering from AD and 1 to images of subject having a cognitively normal status. Thus, if g is close to -1 the subject is more characteristic of AD patient than CN. Otherwise, a g close to 1 is more characteristic of CN subject.

In opposite to the patch-based method proposed by ([Coupé et al., 2012b](#)) that provides a grading value for each considered voxel, for each subject sparse-based grading provides a single value representing the global similarity of the brain structures.

3.1.3 Multiple instance learning

Instead of considering the similarity of each voxel, the multiple instance learning method proposed by (Tong et al., 2014) considers the intra-subject’s patch variability in a graph-based approach. In this method, each image is considered as a bag of cubic fixed-size patches, Because the set of patches is equal to the number of voxel representing the brain, the number of patches is extremely high. Thus, only a subset of the most informative patches are extracted. To extract the most informative patches, probability values that represent the discriminative power of each voxel are estimated with a elastic net regression. Afterward, an iterative step is used to avoid the overlap between selected patches. This step starts with the selection of the voxel having the highest probability, next the probabilities of its neighborhood within a predefined spatial distance are set to 0. Then, the voxel with the second highest probability is selected and the same process is carried out until K patches are extracted.

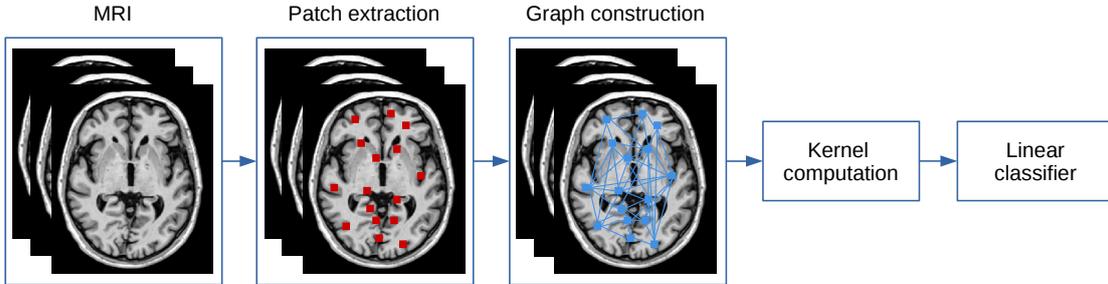


Figure 3.1: Illustration of the multiple instance learning graph, from left to right: The patches are extracted from the most discriminant voxel with a spatial constraint. Next, the graph is built using all the patches extracted. A graph kernel is computed afterwards and used as input of a linear classifier method.

As described in (Tong et al., 2014), the method proposed to integrate the relationship among patches extracted. The main assessments is that such relationship can provide complementary information. Therefore, a graph is computed for each image as follows:

N bags composed of K patches are extracted in the previous step. The intensity values within each patch are organized into a feature vector p such as $B_i = \{p_{i,1}, \dots, p_{i,j}, \dots, p_{i,K}\}$. Finally, the training data is set as $\{(B_1, y_1), \dots, (B_i, y_i), \dots, (B_N, y_N)\}$, with $y_i \in \{0, 1\}$ is related to the pathological status of the bag B_i . In this approach the nodes of the considered graph G_i are the patches of the bag B_i . The distances between each node are defined in the matrix W^i defined with a L_2 norm as follows:

$$W_{a,u}^i = \|p_{i,a} - p_{i,u}\|_2^2 \quad (3.6)$$

where $p_{i,a}$ and $p_{i,u}$ are two different patches from the same bag B_i . W^i represents the relationship between the patches of the bag B_i . In the same way that it has been

done in (Zhou et al., 2009), who captures similarities among graphs with a linear learning method trained using a kernel function K_G is defined. The kernel function K_G is computed as follows:

$$K_G(B_i, B_j) = \frac{\sum_{a=1}^K \sum_{b=1}^K d_{i,a} d_{j,b} k(p_{i,a}, p_{j,b})}{\sum_{a=1}^K d_{i,a} \sum_{b=1}^K d_{j,b}} \quad (3.7)$$

where $d_{i,a} = 1 / \sum_u W_{a,u}^i$, $d_{j,b} = 1 / \sum_v W_{b,v}^j$, the kernel function $k(p_{i,a}, p_{j,b}) = \exp(-\gamma \|p_{i,a} - p_{j,b}\|_2^2)$, $p_{i,a}$ and $p_{j,b}$ are the patches coming from the set B_i and B_j , respectively.

On the opposite of the two previous methods presented that capture inter-subjects' similarities. This last patch-based method capture the intra-individual variability by comparing the most discriminative patch. This method aims to capture specific pattern according to the pathological status.

3.2 Patch extraction

In the original method proposed by Coupé et al. (2011, 2012b) the patch extraction process is composed of two steps. First, a template selection is carried out. The most similar images from the training library are selected with a global image similarity measure. To accelerate the patch extraction step, a first selection is conducted with a global similarity measure. Thus, a L_2 norm is computed for each images of the training library as follows:

$$d(X, Y_t) = \|X - Y_t\|_2^2, \quad (3.8)$$

where X is the image under study and Y_t is the image from the training library. Finally, only the $N\%$ of the most similar images (*i.e.*, images having the smallest distances with X) are selected for the patch extraction process. Second, the patches candidates are extracted from a search window around the coordinate of the voxel under study. As it has been proposed in Coupé et al. (2008), a patch preselection is also carried out to reduce the computational time. The original approach combines an information based on a luminance and a contrast criterion to achieve the patch preselection. The contrast criterion is based on the structural similarity measure (SSIM) (Wang et al., 2004). Finally, the preselection measure can be expressed as follows:

$$pp = \frac{(2\mu_i \mu_{t,j})}{(\mu_i^2 + \mu_{t,j}^2)} \frac{(2\sigma_i \sigma_{t,j})}{(\sigma_i^2 + \sigma_{t,j}^2)} \quad (3.9)$$

where μ and σ are respectively the mean intensity and the standard deviation of the patch under consideration. i denotes the patch from the image X and j, t denotes the patch from the training image Y_t .

Thus, only patches having a pp value superior to a certain threshold are considered into the grading process. This results in a grading of the hippocampus structure

around few minutes. The major strength of this method is that the patch extraction process does not need a non-rigid registration step. Indeed, the only use of affine transformation enables to extract good patch correspondences. However, several limitations appear in this technique. First, the use of global image similarity is computationally expensive and the information redundancy is high since a large overlap between patches extracted can occur. Second, even though the computational time is good compared to the previous method described in the Chapter 2, with the introduction of a large data library, the need of a more effective patch extraction approach is necessary.

3.3 Optimized patch-match

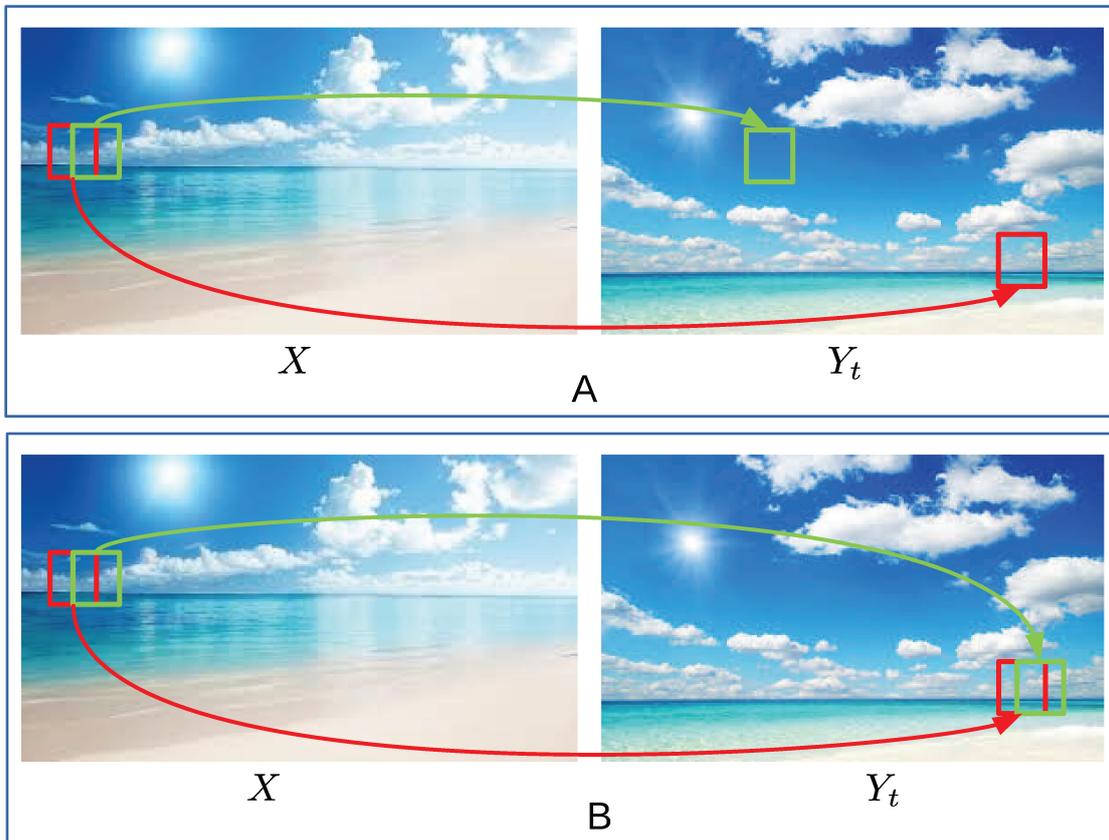


Figure 3.2: Illustration of the propagation step of the patchmatch procedure [Barnes et al. \(2009\)](#). A: a random initialization assigns a patch in the image Y_t to the green patch from the image X . B: the good correspondence of the red patch in the image X is used to propagate the information to the green patch.

In this context, the patchmatch algorithm has been introduced for fast editing of natural images ([Barnes et al., 2009](#)). The approach is based on a semi-random

iterative approximate nearest neighbor (ANN) search. The method consists in randomly associating each patch of the image X with the patches in the image Y_t . Then, an iterative refinement process is carried out, based on the propagation of good matches between adjacent patches, and on the random selection of candidates. The combination of these two steps offers in practice a very fast convergence, and since no preprocessing is performed, the complexity of the method only depends on the number of pixels to be processed in the image X . The key idea of this method is that, based on the spatial coherence of patches, good matches can be propagated to spatially adjacent patches in the image under analysis (see Figure 3.2).

This approach has been optimized for three-dimensional images by Giraud et al. (2016). The method is based on three steps: a constrained initialization, a propagation step, and a random search in the training library (see Figure 3.3).

The constrained initialization consists of the random association of each patch from the image X with a patch of the image Y_t . To ensure, the fast convergence and avoid the correspondence of patch with low intensity distance but with a high spatial distance, the random initialization is carried out in a window search around the coordinate of the voxel of X under study.

Then, in order to improve the initial map of correspondence M , a propagation process and a random search are iteratively carried out. The propagation step is the key step of the patchmatch algorithm. It consists in using the spatial coherency of the patches. The process considers the adjacent patches in terms of 6-connected voxels (*i.e.*, patches at the coordinate: $(x \pm 1, y, z)$; $(x, y \pm 1, z)$; $(x, y, z \pm 1)$). With respect to the position of the voxel under study, the process use the patch adjacent patch in M . The propagation is validated if the intensity distances are lower than the current candidate patch. Finally, after the propagation step, a random search in a fixed size window is carried out. The use of random search after the propagation enables to find a better match is the most-likely good candidate area with avoiding possible local minimal.

The application of the optimized patchmatch algorithm has significantly improved the computational time of the patch-based segmentation. For example, the segmentation of the hippocampus is carried out in less than one second.

3.4 Influence of the parameters

In our works, we implemented the optimized patchmatch developed by Giraud et al. (2016) in order to perform PBG (Hett et al., 2016). To ensure the best performance of patch-based grading using the optimized patchmatch, a few parameters have to be optimized. Thus, in this section, we are studying the influence of the number of extracted patches, the influence of the patch size and finally the influence of the number of template images used for computing grading values. Since the CN and AD group are less prone to noise than the MCI group due to the heterogeneity of the data, we evaluate the different parameters with the classification accuracy of

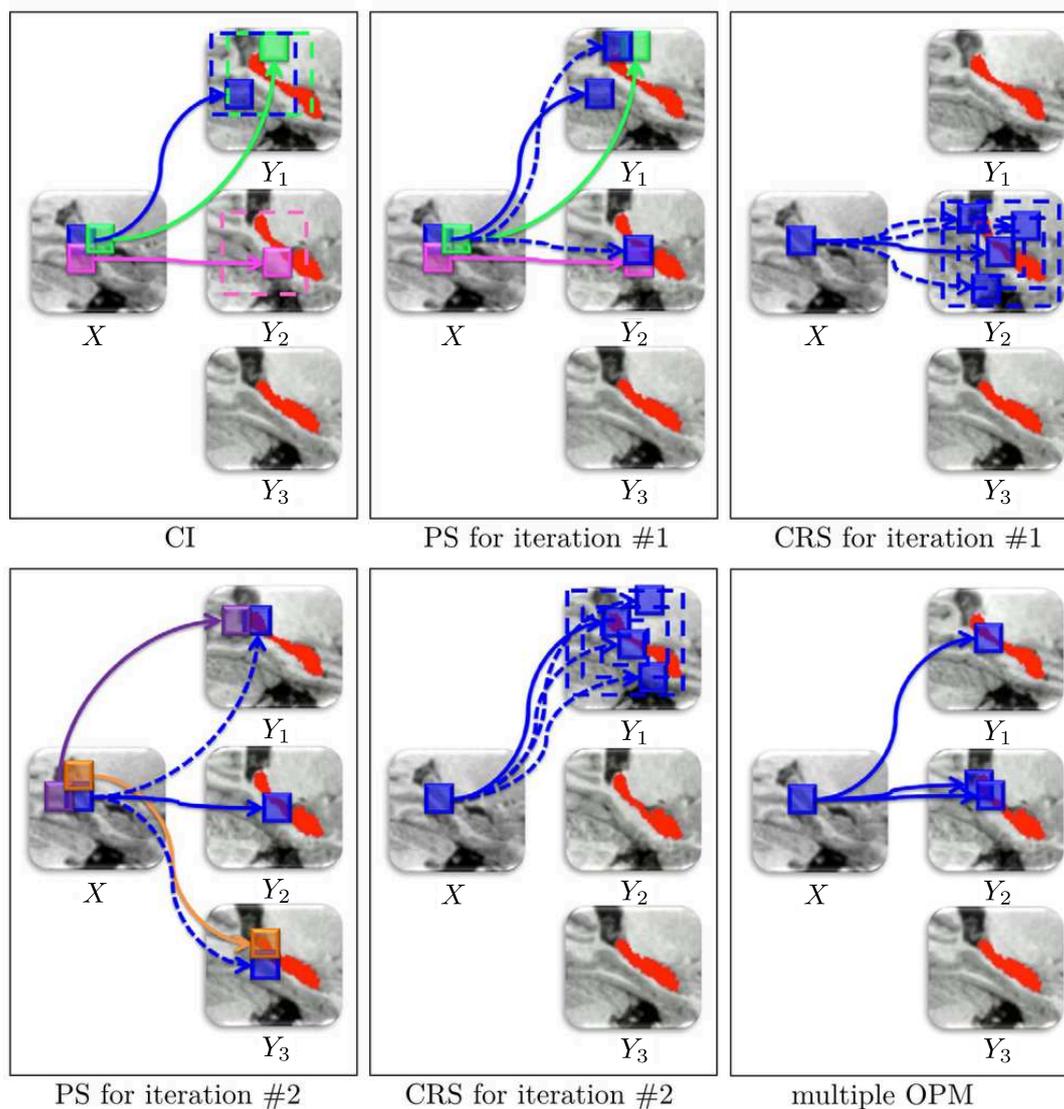


Figure 3.3: Steps of the optimized patchmatch approach. The steps are illustrated for the blue patch in S . The green, pink, purple and orange patches represent the adjacent neighbors of the blue patch. During the constrained initialization (CI), the patches of the subject S are associated (solid lines) to a random patch of the library in a search window depending on their position in S . The propagation step is represented for the first and second iterations. Offset matches of adjacent patches already processed during iteration are tested to improve blue patch matching (dashed lines). The constrained random search is also represented for the first and second iterations. Random matching tests are performed in a search box. *Source: Giraud et al. (2016)*

images from CN and AD patients acquired at the baseline. We split our database into two datasets, one is composed of training images and the second is composed of

Table 3.1: Description of the data used for the evaluation of the optimal parameters for the optimized patch-based grading. All data are from ADNI1 dataset.

Description	Training		Testing	
	CN	AD	CN	AD
Number of subjects	226	186	50	50
Ages	76.5 ± 5.0	75.3 ± 7.4	75.4 ± 5.2	74.5 ± 8.2
Sex	117/109	98/88	28/22	30/20
MMSE	29.05 ± 0.9	22.8 ± 2.9	29.3 ± 0.8	23.1 ± 2.3

testing images (see Table 3.1). Testing images are randomly chosen into the ADNI1 dataset. Moreover, the testing dataset is perfectly balanced to avoid any bias due to the different size of groups. Finally, we evaluate the performance of our patch-based grading feature with the area under the curve (AUC) measure.

3.4.1 Number of patches

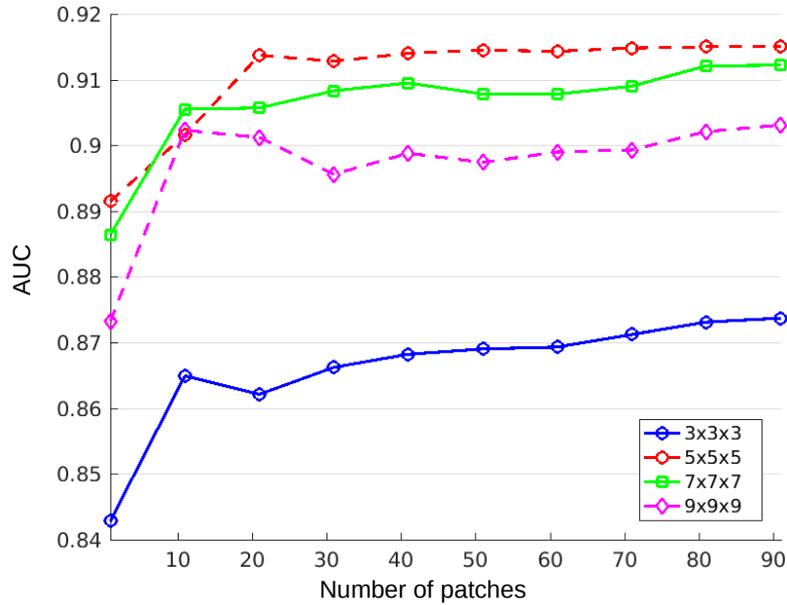


Figure 3.4: The curves represent the classification performance in terms of AUC for CN versus AD classification. Four size of patches have been compared, the results show that patch-based grading method obtains its best performances when around 50 patches at extracted at the size of $5 \times 5 \times 5$ voxels.

First, we evaluate the impact of the number of patches extracted and its correlation with the patch size. The experiments the number of patches from 1 to 100 combined with the size of patch from $3 \times 3 \times 3$ to $9 \times 9 \times 9$ voxels (see Figure 3.4). In

our experiments, we did not constraint the number of patch extract from the images according to their pathological status (*i.e.*, it is possible that $N - 1$ patches are extracted from AD patients and only 1 patches is extracted from the CN subjects). The results show that for all size of patches, there is an improvement of the classification performances with the increase in the number of patches. However, beyond 50 patches, the classification performances do not increase significantly while the computational times increase significantly. It takes around 2 seconds to compute the grading values over the hippocampus with 50 patches when it takes almost 10 seconds to do the same thing with 100 patches. Figure 3.5 represents grading maps estimated with different number of patches. We can notice that only one patch take an arbitrary decision about the status of the tissues at each voxel. With the increase of patch number, the grading maps become smoother, indicating that grading values are better estimated.

Finally, the best results have been obtained with patches of $5 \times 5 \times 5$ voxels. Indeed, the patches of $3 \times 3 \times 3$ voxels are too small to correctly capture the local context while large patches such as $7 \times 7 \times 7$ and $9 \times 9 \times 9$ voxels might suffer from the lack of diversity of the training data. In other words, the complexity of information encoded increases with the size of the patch. Therefore, it becomes necessary to have a more extensive training database.

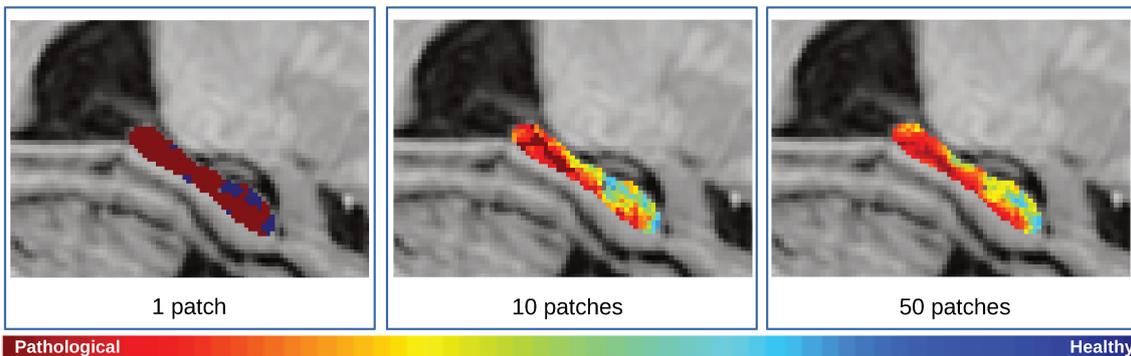


Figure 3.5: This illustration represents grading maps estimated on MRI from an AD patient. Grading maps are computed with different number of patch extracted, from 1 to 50 patches.

3.4.2 Size of the training library

Second, to assess the impact of the number of template images in the grading maps estimation, a comparison of the performances has been conducted for different size of patches with different size of template library (see Figure 3.6). For each size of patches, grading values have been computed with 1 template images in each group (*i.e.*, CN and AD), to 120 template images. The templates images have been ordered with respect to the global distance computed with the sum of square distance

(SSD) of the image under analysis and the images from the training database. In this experiment with select the n closest template images. In this experiment, the optimized patch-match method is set up with the optimal number of patches found in the previous section (see Section 3.4.1).

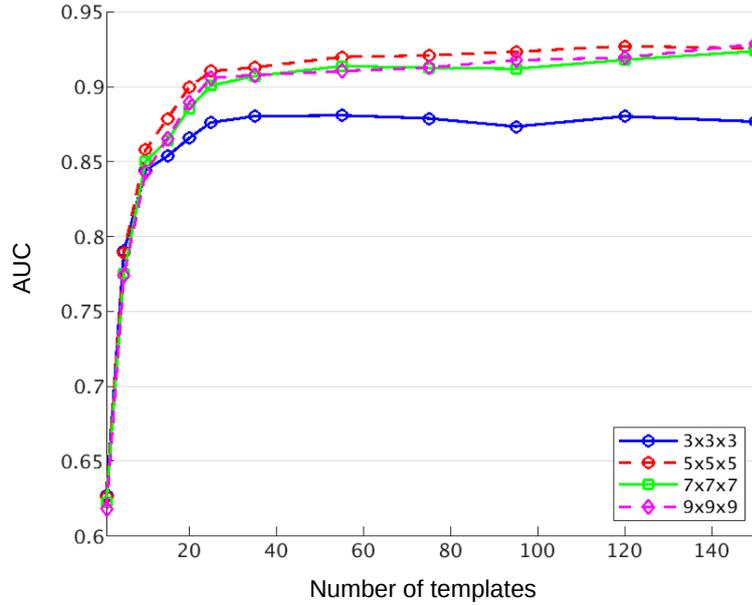


Figure 3.6: The curves represent the classification performance in terms of AUC for CN versus AD classification. The size of templates library has been evaluated from 1 template to 120 templates for each group (*i.e.*, CN and AD).

The experiments show that for any size of patch, the use of the most similar template images from CN and AD library does not enable to capture the alteration related to the disease. To obtain satisfying results compared to the original patch-based grading method, a minimum of 30 templates in each training library must be selected. The best results are obtained with the patch size of $5 \times 5 \times 5$ and 60 template images in each pathological group that compose the training library. Indeed, as illustrated in the Figure 3.7, the grading values are poorly estimated with the use of one template image for each pathological status, even though they are the images templates with the smallest global distance. The increase of the number of template images leads to a refinement of the estimation that results in an improvement of classification performances.

As illustrated in the previous section, patches $3 \times 3 \times 3$ voxels are too small to efficiently capture structural pattern. The best performances are reached with 30 template images for each training libraries. The smallest number of templates needed for this size of the patch could be explained by the higher number of similar area in the templates since the areas to compare are small. On the opposite, large patches need more template images to reach best performances. We can notice

3. Patch-based grading

that patch sizes of $7 \times 7 \times 7$ and $9 \times 9 \times 9$ voxels obtain their best classification performances around 100 template images in each training libraries. With a greater number of template images, the classification performances of large patches are similar to patches of $5 \times 5 \times 5$. It might be reasonable to assume that large patches could obtain slightly better results with an even bigger number of template images. However, their computational times are expensive compared to the smaller patches. Therefore, we decide to set up the optimized patch-match algorithm with the patch size of $5 \times 5 \times 5$ voxels with 60 templates for each pathological status.

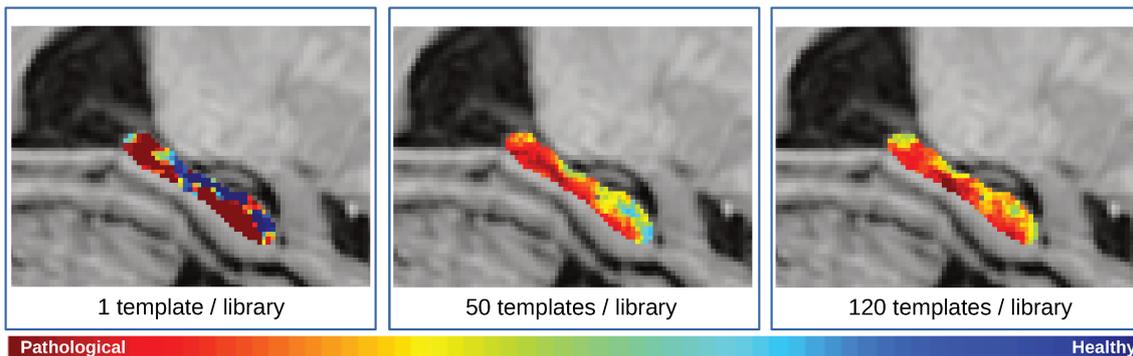


Figure 3.7: This illustration represents grading maps estimated on MRI from an AD patient. Grading maps are computed with different number of patch extracted, from 1 to 50 patches. The grading values are computed with patches of $5 \times 5 \times 5$ voxels.

3.5 Conclusion

In this chapter, we present three different patch-based methods from the original patch-based grading framework to the most recent proposed methods. Although the recently proposed methods involved advanced approaches such as sparse representation or graph-based modeling, they have not significantly improved classification performances compared with the original patch-based grading framework based on an exponential kernel. Indeed, compared on the same brain structure (*i.e.*, see Table 2.1), the SNIPE method obtain best performances for AD detection and prediction with 88% and 71% of accuracy within the hippocampus, respectively. Moreover, with the recent development of fast patch extraction provided by the optimized patch-match, the patch-based grading feature can be estimated in less than a second.

Part conclusion

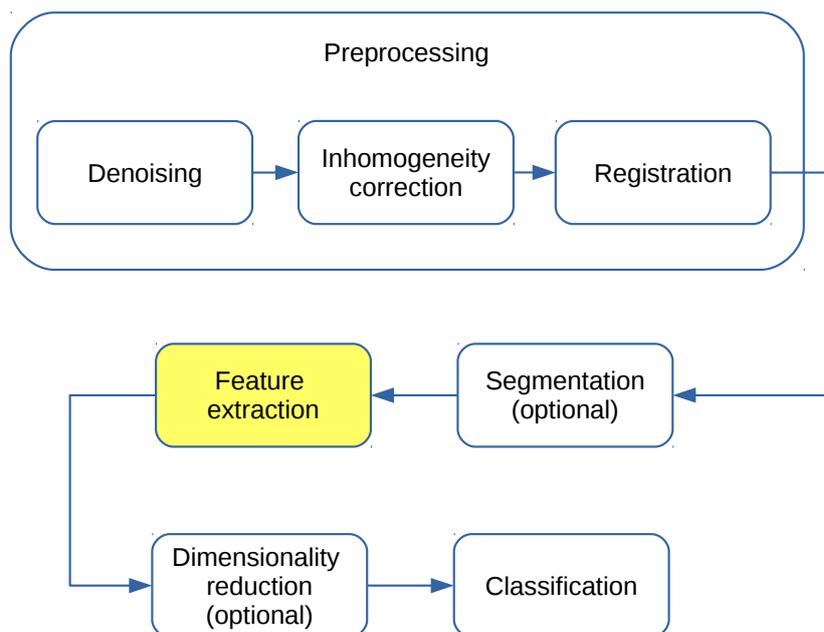


Figure 3.8: As highlighted in this schema, during this thesis, we have focused on the development of new imaging-based features to improve the detection and prediction of Alzheimer’s disease.

Motivations

As illustrated in Figure 3.8, in this thesis, we focused on the development of new efficient biomarkers for the early detection of AD. We have mainly been interested in the improvement of patch-based grading framework. Indeed, the introduction of patch-based methods has led to a better estimation of structural alterations into specific structures such as hippocampus compare to previous methods. Some elements tend to show that a better analysis of hippocampus could increase the performance for the early detection of AD.

Indeed, on the one hand, a large proportion of data comes from different neuroimaging campaign and has been acquired from different neuroimaging centers. Therefore, even though an acquisition protocol has been defined, the images are

acquired with different imaging devices. This results in a variability of MRI intensities that might not be corrected by the preprocessing step. It seems essential to develop a method robust to acquisition variability. Also, improving patch comparison by highlighting discriminant signal into a patch could lead to an improvement of the performances for early detection of AD. Indeed, a part of the signal embedded in a patch can contain all discriminant information. For instance, edge and texture signals can be discriminant while the other part of the signal can be not informative since it may just add noises within each pathological group since it can only be correlated to other factors.

On the other hand, structural MRI (s-MRI) is not the only MRI modality usable within the PBG framework, for example, diffusion MRI (d-MRI) can also be integrated into patch-based grading study. Studies based on d-MRI usually capture axonal alterations by analyzing modifications of white matter with tractography approaches. However, d-MRI can also be used to capture micro-structural alteration of grey matter structures. The application of PBG method on this MRI modality can improve early detection performance since such alterations are considered to occur before structural alteration. A limitation of d-MRI comes from its native resolution that is often lower than s-MRI and reduces the ability of d-MRI for providing good biomarkers for AD detection. However, recent super-resolution methods demonstrate good results and have addressed the issue of low resolution in many pattern recognition problems.

Besides, the hippocampus is a heterogeneous structure. Indeed, the hippocampus is composed of different subfields having distinct characteristics. Moreover, during the last decade, postmortem and animal-based studies have shown that the hippocampal subfields are not equally impacted by AD. Indeed, these studies showed that some subfields suffer from stronger atrophy at the late stage of AD than others. Therefore, a fine-grained analysis of structural hippocampus alterations at subfields scale could also lead to an improvement for the early detection of AD.

Finally, although state-of-the-art methods for the early detection of AD have mostly focused on a robust extraction of inter-subjects' similarities in specific ROIs or different scale of analysis (*i.e.*, voxel, patch, etc), some works developed models that capture intra-subjects' variability of alterations. These methods have demonstrated that an efficient representation of the topology of alterations by modeling the relationship of structural changes between the different brain structures can help to improve AD detection and prediction. Indeed, it seems that the relationships of the structural alterations between brain structures form a specific pattern that provides relevant information on the progression of the dementia. Therefore, modeling this pattern could improve the early detection of AD.

All these elements raise questions that suggest four ways of research to improve PBG methods.

- How to improve patch comparison to enhance informative signals encoded the patches?

- How to develop an efficient multimodal approach involving d-MRI to improve performances for early detection of AD?
- Is the study of hippocampal alterations at a finer scale, such as hippocampal subfield analysis, effective for early detection of AD?
- Does the combination of inter-subject similarities and intra-subject variability make it possible to increase AD prediction performance?

Therefore, in this thesis we will explore these four approaches as summarized in the followings.

Research contributions and thesis outline

The contribution of this thesis are summarized below.

Adaptive fusion of texture-based grading

Computer-aided diagnosis system are usually based on method using intensity or grey matter maps. However, it has been shown that texture filters improve classification performance in many cases. The aim of this work is to improve performances of patch-based grading framework with the development of a novel texture-based grading method. In this work, we study the potential of multi-directional texture maps extracted with 3D Gabor filters to improve patch-based grading method. We also propose a novel patch-based fusion scheme to efficiently combine multiple grading maps. To validate our approach, we study the optimal set of filters and compare the proposed method with different fusion schemes. In addition, we also compare our new texture-based grading biomarker with state-of-the-art methods. This first contribution dedicated to texture-based grading will be detailed in Chapter 5.

Multimodal patch-based grading

Our next contribution is the application of PBG framework to d-MRI modality. In this work, we propose to use a similar framework on s-MRI and d-MRI. We compare our new grading-based d-MRI features with basic MRI and d-MRI biomarkers. In addition, we propose a new framework to efficiently merge PBG maps from s-MRI and d-MRI. First, although d-MRI performance lower results than s-MRI in the late stage of AD, we demonstrate that the proposed d-MRI biomarkers obtain competitive results for the early detection of AD. Finally, the experiments conducted have shown that our multimodal biomarker merging PBG from s-MRI and d-MRI obtain robust classification performances. Indeed, our multimodal PBG obtains best classification performances for late and early detection of AD. This second contribution dedicated to multimodal grading will be detailed in chapter 6.

Hippocampal subfields grading

To date, the patch-based grading framework provides competitive hippocampal biomarker. However, this structure is complex since the hippocampus is divided into several heterogeneous subfields not equally impacted by AD. Former *in-vivo* imaging studies only investigated structural alterations of these subfields using volumetric measurements and microstructural modifications with mean diffusivity measurements. The aim of our work is to study the efficiency of hippocampal subfields compared to the whole hippocampus structure with a multimodal patch-based framework that enables to capture subtler structural and microstructural alterations. To this end, we analyze the significance of the different hippocampal subfields for AD diagnosis and prognosis with volumetric, diffusivity measurements and a novel multimodal patch-based grading framework that combines structural and diffusion MRI. The experiments conducted in this work showed that the whole hippocampus provides the most discriminant biomarkers for advanced AD detection while biomarkers applied into subiculum obtain the best results for AD prediction. This third contribution dedicated to hippocampal subfields grading will be detailed in chapter 7.

Graph of brain structures grading

Nowadays, numerous studies have proposed biomarkers to perform early detection of AD. Some of them have proposed methods based on inter-subject similarity while other approaches have investigated framework using intra-subject variability. Therefore, our last contribution is the development of a graph-based model that embeds inter-subjects' similarity and intra-subjects' variability information. Indeed, in this work, we propose a novel framework combining both approaches within an efficient graph of structures grading. First, we apply this framework with the definition of hippocampal subfields studied previously. We demonstrate that our novel approach obtains competitive results compared to former methods based on the hippocampus. Second, since best methods are based on an whole brain analysis, we apply our novel graph-based approach within the whole brain. This results in state-of-the-art performance for the prediction of subject conversion to AD. Finally, we propose an unified framework to combine whole brain structures representation with the hippocampal subfields. The experiments conducted in this last work demonstrate state-of-the-art classification performances for AD detection and prediction. This last contribution dedicated to graph of structure grading will be detailed from the chapter 8 to 11

Part II:

Advanced hippocampus biomarkers

The first contributions of this Ph.D are based on the design of advanced hippocampus biomarkers. In this part, we will present a novel adaptive framework to fuse multiple grading maps. This adaptive fusion will be used for two different applications. The first application is a new texture-based grading method that use texture filters to improve patch comparison. The second one is a multimodal patch-based grading method that involves s-MRI and d-MRI. This multimodal approach aims to improve AD performance detection by combining structural and microstructural information. Finally, we we will present a multimodal study to measure the hippocampal subfields effectiveness for the early detection of AD.

Introduction

To date, patch-based grading has been applied only to structural MRI intensities. However, some recent works have proposed to apply patch-based segmentation methods by using different inputs (*i.e.*, modalities). Indeed, recently segmentation methods based on patch-based framework have been improved with the use of multi-source approaches (*i.e.*, image derivatives such as gradient).

However, so far, a limitation occurs in the fusion methods designed to combine information coming from the different image sources. On the one hand, a recent multi-features patch-based segmentation method has proposed to fuse the different estimator maps by a straightforward average. On the other hand, a multi-contrast patch extraction method has proposed to compute the similarity of patches from two different sources with a multi-contrast norm. Such methods assume that all features or modalities have the same relevance. However, this is not always true since there is in some cases similar patches in the template library. Therefore, in this part, we propose a new adaptive fusion scheme that provides a locally adaptive criterion for weighting grading values from each MRI sources at each voxel.

We apply this adaptive fusion scheme to develop a novel texture-based grading method. This novel method takes advantage of the adaptive fusion of directional derivative filters. The use of these filters aims to enhance informative signal in different directions. This approach obtains state-of-the-art results for AD detection compare to most advanced methods.

Besides structural MRI, other MRI modalities such as diffusion MRI could provide relevant information about brain alterations. Such modality can also be used as input of path-based grading methods. Consequently, we investigate patch-based grading method using diffusion MRI modality. The comparison of PBG based on these two different MRI modalities demonstrates that structural MRI provides the best results for AD detection. However, the application of diffusion MRI as input of PBG enables to improve results for AD prediction. Consequently, we apply our novel adaptive fusion scheme to fuse patch-based grading based on both structural and diffusion MRI. The experiments show that the multimodal patch-based grading based on these two modalities obtains competitive results for AD detection and prediction, confirming the interest of using structural and diffusion MRI.

Finally, we propose to study the hippocampus structure at a finer grained scale with the analysis of the hippocampal subfields. Therefore, we carry out a comparative multimodal analysis to emphasize the advantage of applying our method into specific hippocampal subfields. The results of our experiments based on volumetric measurement are in line with previous hippocampal subfields studies for the most discriminant atrophy at late stages of AD. Moreover, we also show that subiculum provides good multimodal biomarker at the early stage of AD and enables to improve prediction performances compared to the use of the whole hippocampus.

Key-words

Patch-Based Grading, Multi-Source, Multi-Modal, Multi-Feature, Adaptive Fusion, Structural MRI, Diffusion MRI, Texture-Based Analysis, Hippocampal Subfields

Contents

- The chapter 4 presents state-of-art approaches that propose multimodal, and multi-texture fusion methods. In this chapter, we propose an adaptive fusion scheme that attempts to fuse efficiently several patch-based grading maps.
- In the chapter 5, we apply our new adaptive fusion framework with multiple directional texture filter and propose a texture-based grading method. In this chapter, the three main contributions are the applications of a bank of Gabor's filters that provides directional information of the texture, the application of our novel adaptive fusion scheme and the aggregation of final fused grading values over the hippocampus with a histogram representation. We validate our approach with a comparison of different fusion schemes, different texture filters. Finally, we investigate the optimal set of direction textures.
- The chapter 6 introduces the DTI patch-based grading and a new multimodal patch-based grading that combines structural MRI and diffusion MRI. In this chapter, we apply our novel fusion scheme to fuse mean diffusivity grading and structural grading maps. This produces more robust features obtaining state-of-the-art results for detection and prediction of Alzheimer's disease.
- The chapter 7 studies the alterations of hippocampal subfields with our proposed multimodal patch-based grading. In this chapter, we propose a comparative study based on a multimodal framework to analyze changes that occur into the hippocampal subfields. This work aims to validate whether a fine-grained analysis of hippocampus based on its subfields can provide more efficient biomarkers and identify the best hippocampal subfields for AD detection and prediction.

Related publications

The works of this part are based on the following papers:

Peer-reviewed international journals

- [1] Multimodal hippocampal subfield grading for Alzheimer's disease classification. K. Hett, V.-T. Ta, G. Catheline, T. Tourdias, J. V. Manjón, and P. Coupé. Scientific Reports (Sci. Rep), submitted
- [2] Adaptive fusion of texture-based grading for Alzheimer's disease classification. K. Hett, V.-T. Ta, J. V. Manjón, and P. Coupé. Computerized Medical Imaging and Graphics (COMPUT MED IMAG GRAP)

Peer-reviewed international conferences

- [3] Adaptive fusion of texture-based grading: Application to Alzheimer's disease detection. K. Hett, V.-T. Ta, J. V. Manjón, and P. Coupé. Patch-based Techniques in Medical Imaging (PatchMI@MICCAI), 2017
- [4] Patch-based DTI grading: Application to Alzheimer's disease classification. K. Hett, V.-T. Ta, R. Giraud, M. Mondino, J. V. Manjón, and P. Coupé. Patch-based Techniques in Medical Imaging (PatchMI@MICCAI), 2016

Chapter 4

Adaptive fusion

4.1	Introduction	57
4.2	Background	57
	4.2.1 Multi-modal fusion for Alzheimer’s disease classification	57
	4.2.2 Multi-spectral/features fusion of patch-based methods	59
4.3	Method	61
4.4	Conclusion	62

4.1 Introduction

Since the introduction of patch-based framework for brain image analysis (Coupé et al., 2011), numerous improvements have been proposed to enhance the performance of this framework. Some methods have proposed to replace the exponential kernel by a sparse-based approach based on dictionary learning to estimate the most discriminant patches extracted from the training library (Liu et al., 2012; Tong et al., 2013, 2017a). However, despite an increased complexity of these patch-based methods, their performances have not demonstrated significant improvement compared to the methods based on an exponential kernel.

Recently, a patch-match optimization technique (Barnes et al., 2009) was adapted to the segmentation of 3D structural MRI (Giraud et al., 2016). This method drastically reduces the computation time of the patch extraction step. Indeed, the use of this optimization method leads to a computation of the hippocampus segmentation under a second with state-of-the-art precision. This improvement has enabled the use of multi-source inputs. The use of multi-source method requires fusion of the different sources. This fusion step can occur at different stages of the patch-based pipeline. We can cite two different fusion schemes. The first is a late fusion of the different features computed by the method for each source, while the second is an early fusion of the features produced at the patch extraction step. This last group of methods demonstrates an improvement of patch-based method performances. However, a major limitation occurs since such techniques consider that all estimator maps have the same relevance. Therefore, we develop in this chapter a new adaptive fusion scheme to locally and automatically weight the grading values coming from each sources.

4.2 Background

This section is divided into the description of multi-modal fusion scheme proposed in the literature that can be applied for any kind of method, and fusion approach specially dedicated to the patch-based framework.

4.2.1 Multi-modal fusion for Alzheimer’s disease classification

To date, several works proposed to improve AD classification using the complementarity of different medical imaging modalities. Thus, some methods proposed to fuse features based on MRI with biomarkers based on PET, DTI, fMRI. Usually, the proposed multimodal approaches extract features from each modality separately and combine them into methods based on several classification layers.

Figure 4.1 illustrates a generic pipeline usually applied to multimodal image analysis. This first group of multimodal approaches proposed to extract features such as

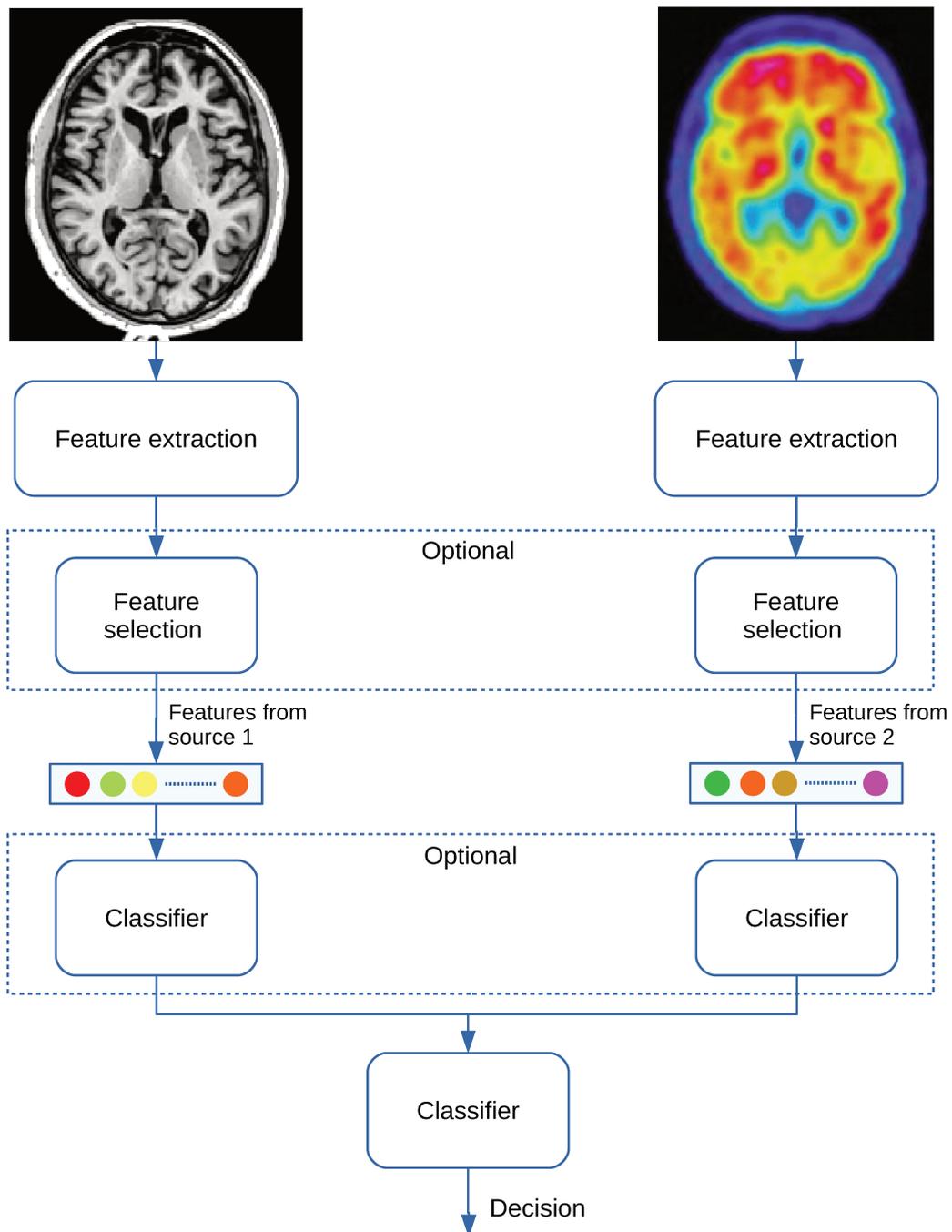


Figure 4.1: This figure illustrates the general framework usually applied in multimodal image analysis method. Features are extracted independently from each considered modality. Afterward, an optional step is used to reduce the dimensionality via a feature selection. The features selected are individually used as input of an intermediate layer of classifiers or concatenate into a single vector. Finally, a final classifier is used to make the final decision.

cortical thickness, grey matter intensity or PET intensities... This group of methods computes a concatenation of the different feature vectors into an extended vector of features to use it as an input of classifier method. The classification methods used are often designed to deal with heterogeneous data such as multi-kernel learning approaches or random forest (Kohannim et al., 2010; Zhang et al., 2011; Gray et al., 2013). Other methods proposed to reduce the dimensionality of the features by applying a feature selection or a linear transformation into a lower dimensionality space (Hinrichs et al., 2011; Dyrba et al., 2015). Recently, more advanced methods proposed to use a stack of deep learning networks, each network having to handle data from different modality (Calhoun and Sui, 2016; Shi et al., 2018), for instance, a method proposed to find a latent hierarchical patch representation with a deep Boltzmann method (Suk et al., 2014). Finally, another method proposed to integrate brain structure volumes, FDG-PET and CSF biomarkers measures into a non-linear graph fusion (Tong et al., 2017b). These works demonstrated that this kind of framework can improve the classification performances of AD detection and prediction by taking advantage of the complementarity information of each modality.

4.2.2 Multi-spectral/features fusion of patch-based methods

With the development of patch-based methods, some works proposed to fuse images information at a patch scale. Some works proposed an early fusion of features into the patch difference measurement (Kim et al., 2013; Bai et al., 2015), and another proposed to integrate the distance with a multi-contrast semi norm (MSN) distance based on a combination of SSD (Romero et al., 2017), this new metric is defined as follows:

$$MSN_{i,j,s} = \frac{\|P_{x_i} - P_{t_j}\|_2^2 \cdot \|P_{x'_i} - P_{t'_j}\|_2^2}{M(\|P_{x_i} - P_{t_j}\|_2^2 + \|P_{x'_i} - P_{t'_j}\|_2^2)} \quad (4.1)$$

where P_{x_i} represents the patch surrounding the voxel i of the image x , $P_{x'_i}$ is the patch surrounding the voxel j of the image template t . x and t represents the data from the first modality while x' and t' represents the data from the second modality.

These proposed methods show that the use of different input combined with a fusion at a patch scale can lead to better capturing the pattern of the tissue surrounding the voxel under analysis. However, so far, the developed techniques assume that each patch has the same relevance to describe the tissues under analysis as each source has the same relevance. Thus, an optimized way to fuse the patch-based estimators is crucial to deal with the fact that each patch-based estimator can embed signal having unequal similarity relevance.

Taking advantage of the fast patch extraction based on an optimized patch match (OPM) technique, (Giraud et al., 2016) proposed to fuse multiple feature images. This method is based on the combination of images filtered with a gradient filter and raw intensities of structural MRI. Moreover, in his work Giraud et al. (2016) proposed to embed different features with a multi patch size approach. The method proposed to fuse the estimator maps by a straightforward average of the estimators

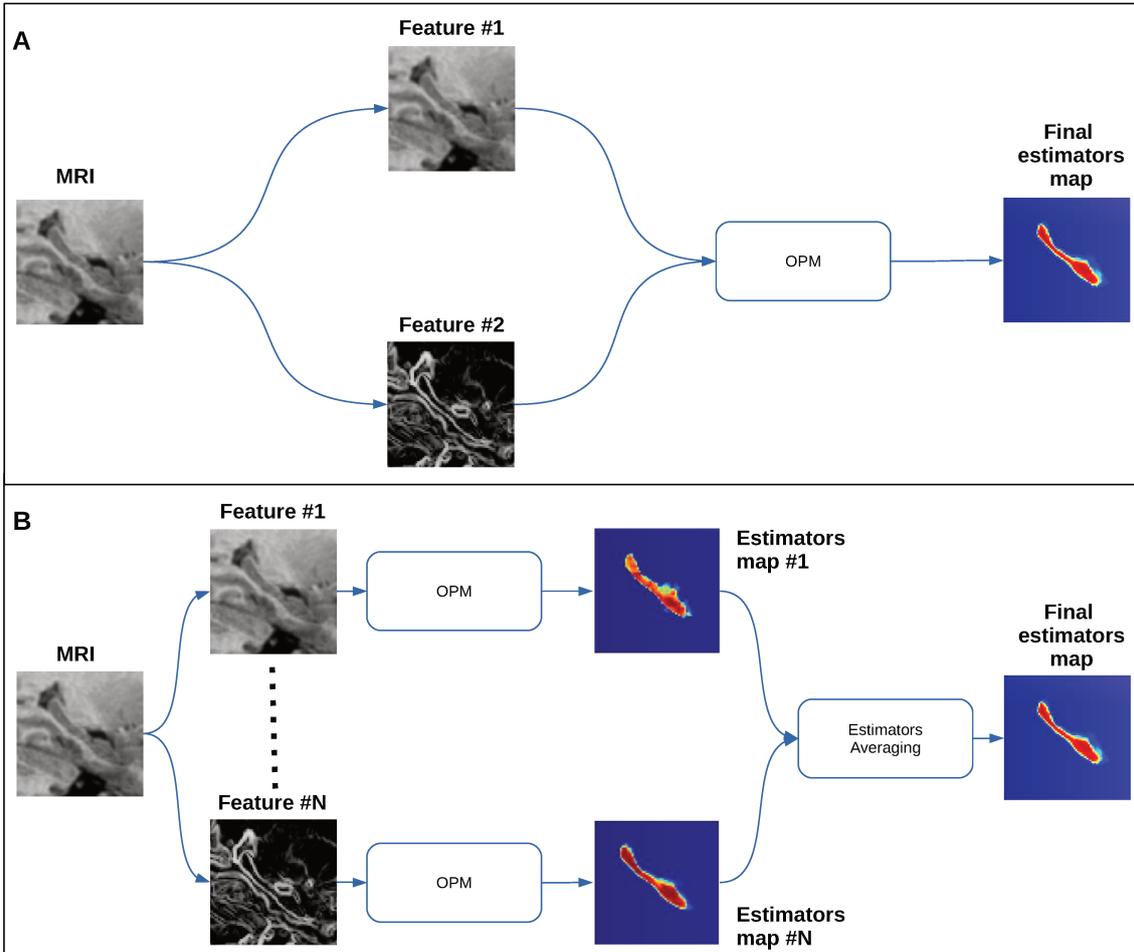


Figure 4.2: Schema representing the early estimator fusion proposed in (Romero et al., 2017) (frame A) and the late estimator fusion proposed in (Giraud et al., 2016) (frame B). The early fusion represented in frame A proposes to extracted patch with a multi-spectral norm that computes the distance of the patches surrounding the voxel under the analysis of two different sources. In opposite, the late fusion of patch-based estimator represented in the frame B is a straightforward average of the estimator computed at each voxel.

at each voxel (see Figure 4.2) as follows:

$$l = \sum_{i=1}^N l^i \quad (4.2)$$

where l is the final map of fused estimators, l^i is the map of estimators from the feature i , and N is the number of features. The optimized patch match technique demonstrated an improvement of the segmentation accuracy compared to the last advanced methods based on dictionary learning and sparse coding approach (Tong et al., 2013; Zhu et al., 2017). Also, this method revealed that texture filters applied on anatomical MRI could improve the performances of patch-based methods.

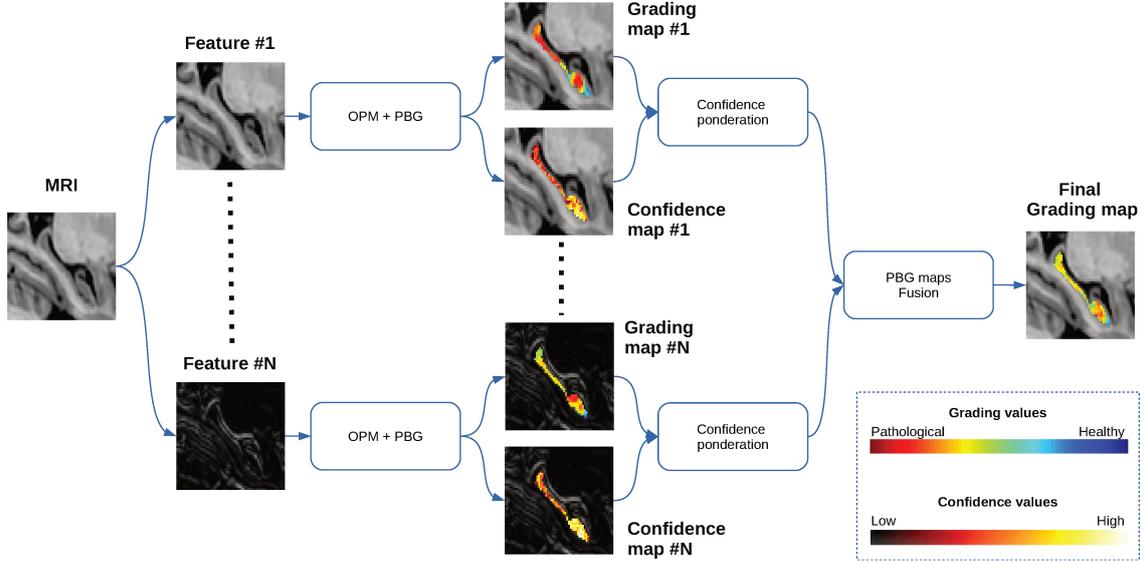


Figure 4.3: Schema representing the pipeline of the proposed adaptive patch-based grading fusion scheme. For each feature, a patch-based grading map is computed. Our technique provides grading maps and confidence maps. Each grading map is weighted with its confidence map and combined afterward with the grading maps from the other sources.

4.3 Method

In this chapter, we propose a novel framework to locally and efficiently fuse multiple grading maps computed from different sources (see Figure 4.3). Our fusion strategy assumes that all the estimated grading maps may not have the same relevance, but more importantly all local grading values g_{x_i} in these maps do not have the same quality. Hence, at each location, we propose to combine grading values derived from different MRI sources according to a confidence criterion. Therefore, the grading value $g_{x_{i,n}}$ at voxel $x_{i,n}$ of a grading map based on the source n , is weighted by:

$$\alpha_{x_{i,n}} = \sum_{t_j \in K_{i,n}} w(P_{x_{i,n}}, P_{t_j,n}) \quad (4.3)$$

that reflects the confidence of grading value $g_{x_{i,n}}$ for source n , here t_j is the voxel j belonging to the training template $t \in T$ and $K_{i,n}$ is the set of similar patches extracted from the training library T at the voxel x_i of the texture map n . Thus, a grading map is computed for each source and provides a grading value at each voxel that is weighted with its degree of confidence $\alpha_{x_{i,n}}$. The final grading value g_{x_i} is given by:

$$g_{x_i} = \frac{\sum_{n \in N} \alpha_{x_{i,n}} g_{x_{i,n}}}{\sum_{n \in N} \alpha_{x_{i,n}}} \quad (4.4)$$

The proposed fusion framework is spatially adaptive and take advantage of having access to a local degree of confidence $\alpha_{x_{i,n}}$ for each grading map n . Basically, the con-

fidence $\alpha_{x_i,n}$ gives more weight to a weak classifier estimated with a well-matched set of patches. This adaptive fusion strategy can be applied to any patch-based processing to combine in the same manner multiple feature maps or different modalities.

4.4 Conclusion

The different fusion schemes proposed in the literature assess that each feature coming from different image sources has the same significance. However, this could be a limiting factor since the relevance of each feature extracted from different sources can be unequal.

Furthermore, the relevance of patch-based grading values are related with the similarities of the patches extracted from the training library. Indeed, patches with low similarity do not encode discriminant information about the structural alterations of subject under analysis brain tissues.

Our method aims to improve patch-based grading estimation by locally adapting the weight given to each weak-classifier from different image sources. Therefore, we proposed a novel adaptive fusion scheme that is based on the similarity of the set of patches extracted with the optimized patch match method ([Giraud et al., 2016](#)). At each voxel a confidence criterion is used to weight the grading value from each image source. In this thesis, we propose two applications of our adaptive fusion with a multi-feature and multi-modality approaches.

Chapter 5

Texture-based grading

5.1	Introduction	64
5.2	Materials	65
5.3	Experiments	65
5.3.1	MRI preprocessing	65
5.3.2	Texture maps estimation	66
5.3.3	Weak classifiers aggregation	67
5.3.4	Implementation	68
5.4	Results	68
5.4.1	Optimal number of directions	68
5.4.2	Comparison grading based on intensity vs. texture	69
5.4.3	Comparison average grading vs. histogram-based grading	70
5.4.4	Comparison of different fusion schemes	71
5.4.5	Comparison of different texture-based filters	72
5.4.6	Comparison to state-of-the-art methods	73
5.5	Discussion	73
5.6	Conclusion	76

5.1 Introduction

In the previous chapter, we described a new adaptive fusion scheme that enables to efficiently merge several patch-based grading maps.

In this chapter, we propose to apply this method to multi-directional derivative texture filters. Indeed, to date patch-based methods have been based on intensities (Coupé et al., 2012b; Tong et al., 2017a) and grey matter probability maps (Liu et al., 2012; Komlagan et al., 2014). However, it has been demonstrated that hippocampus texture improves the detection at early stages of AD (Sørensen et al., 2016b). A VBM method using several textural filter on medial temporal lobe area demonstrated the reliability of texture information for AD detection (Chincarini et al., 2011). Moreover, hippocampus texture enables to improve AD detection compared to hippocampus volume (Sørensen et al., 2016a). This method could potentially capture MRI signal alterations related to neurofibrillary tangles and beta-amyloid plaque deposition, although such alterations are not directly detectable with MRI at current resolution. Besides, a recent study recently showed the efficiency of using edge detection filters to improve of patch-based segmentation (Giraud et al., 2016). This result highlights that patch-based grading methods could be improved by estimating patterns similarity on derivative image features. Therefore, we propose to perform patch-based grading on multiple texture maps obtained with Gabor filters. Gabor filters are designed to detect salient features at specific resolution and direction. These filters were widely used for texture classification (Manjunath and Ma, 1996; Grigorescu et al., 2002; Riaz et al., 2013). The proposed strategy enables to better capture texture modifications occurring at the first stages of the pathology by improving patch comparison.

The first contribution proposed in this chapter is the development of a new texture-based grading framework to better capture structural alterations caused by AD. This new framework proposes multi-directional texture grading based on 3D Gabor filters. Secondly, in order to combine all the grading maps estimated on texture maps, we apply our innovative adaptive patch-based fusion strategy based on local confidence criterion. Moreover, contrary to usual grading-based methods using the average grading values over the considered ROI, we propose a classification step based on a nonparametric grading values distribution representation to better discriminate pathologies stages. In our experiments, we first study the optimal number of Gabor filter directions for AD detection. In addition, we compare different texture filters such as local variance or entropy. We also compare our new adaptive fusion method with different fusion schemes. Finally, to highlight the improvement of classification performances provided by our new framework, we compare our new method with the state-of-the-art approaches and demonstrate its efficiency.

5.2 Materials

Data used in this work were obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) dataset¹. The data used in this study are all the baseline T1-weighted (T1w) MRI of the ADNI1 phase. This dataset includes AD patients, MCI and cognitive normal (CN) subjects. The group of MCI is composed of subjects who have abnormal memory dysfunctions and embed two groups, the first one is composed with patients having stable MCI (sMCI) and the second one is composed with patients with progressive MCI (pMCI), such patients converted to AD during the following 48 months from the baseline (Wolz et al., 2011). The information of the dataset used in our work is summarized in Table 6.1.

Table 5.1: Description of the dataset used in this work. Data are provided by ADNI.

Characteristic / Group	CN	sMCI	pMCI	AD
Number of subjects	226	223	165	186
Ages (years)	76.0 ± 5.0	75.1 ± 7.5	74.5 ± 7.2	75.3 ± 7.4
Sex (M/F)	117/109	150/73	101/64	98/88
MMSE	29.05 ± 0.9	27.1 ± 2.5	26.3 ± 2.0	22.8 ± 2.9

5.3 Experiments

In this section, we describe the different steps of the proposed texture-based grading framework as illustrated in figure 5.1. First, we use multi-directional Gabor filters to extract texture in different directions. Second, a patch-based grading method is applied within each texture map computed. Next, all the texture grading maps are merged with our novel adaptive fusion method. Finally, the final feature considered is a histogram representation of texture grading values in each hippocampus.

5.3.1 MRI preprocessing

All the T1w images were processed using the volBrain system Manjón and Coupé (2016)¹. This system is based on an advanced pipeline providing automatic segmentation of different brain structures from T1w MRI. However, in this work, only hippocampus segmentations were used. The preprocessing is based on: (a) a denoising step with an adaptive non-local means filter Manjón et al. (2010), (b) an affine registration in the MNI space Avants et al. (2011), (c) a correction of the image inhomogeneities Tustison et al. (2010) and (d) an intensity normalization Manjón et al. (2014). Afterwards, MRI were segmented in the MNI space using non-local

¹<http://adni.loni.ucla.edu>

¹<http://volbrain.upv.es>

patch-based multi-atlas methods [Coupé et al. \(2011\)](#). The obtained hippocampus were segmented according to the EADC protocol [Boccardi et al. \(2015\)](#) designed for AD studies.

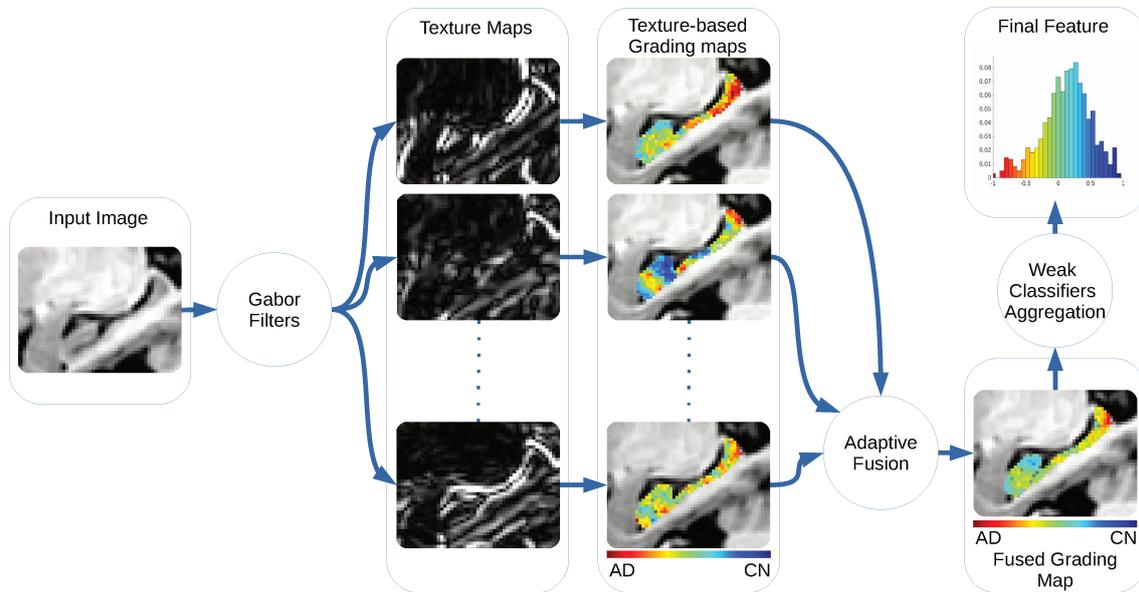


Figure 5.1: Proposed adaptive fusion of texture-based grading framework: from left to right, the T1w input data, the texture maps for different directions, the intermediate texture-based grading maps, the final fused grading map and the histogram-based weak classifiers aggregation.

5.3.2 Texture maps estimation

The estimation of patch similarities could be improved by using texture representation instead of using raw intensities. Indeed, it was demonstrated that the use of edge detectors improves patch-based segmentation accuracy ([Giraud et al., 2016](#)). Moreover, it was also demonstrated that hippocampus textural information plays an important role in AD detection ([Sørensen et al., 2016b](#)). Hence, we propose a new texture-based grading framework that simultaneously captures hippocampus texture alterations and improves patches similarity estimation. In this work, texture information is extracted from MRI using a bank of 3D Gabor filters (see [Figure 5.2](#)). We used Gabor filters since they are designed to detect texture patterns at different scales and directions ([Manjunath and Ma, 1996](#)). Impulse response of the 3D Gabor filter is given by the following equation:

$$h(X, \sigma, F, \theta, \phi) = \frac{1}{(2\pi)^{3/2}\sigma^3} \exp\left(-\frac{x^2 + y^2 + z^2}{2\sigma^2}\right) \hat{h}(X, F, \theta, \phi) \quad (5.1)$$

where $X = (x, y, z)$, σ represents the standard deviation of the Gaussian function. (θ, ϕ) are the orientation angles. F represents the central frequency of the frequency response. Finally, $\hat{h}(X, F, \theta, \phi)$ is given by:

$$\hat{h}(X, F, \theta, \phi) = \exp(j2\pi F(x \sin \theta \cos \phi + y \sin \theta \sin \phi + z \cos \theta)) \quad (5.2)$$

In our method, the texture maps are the magnitude of the signal resulting from the convolution of our Gabor filters bank with the MRI. In the proposed pipeline (see Figure 5.1), the preprocessed MRI of the subject under study is filtered with a bank of Gabor filters to obtain multiple texture maps. All the training library is also filtered with the same filters bank. Therefore, for each texture map, a texture-based grading map can be estimated.

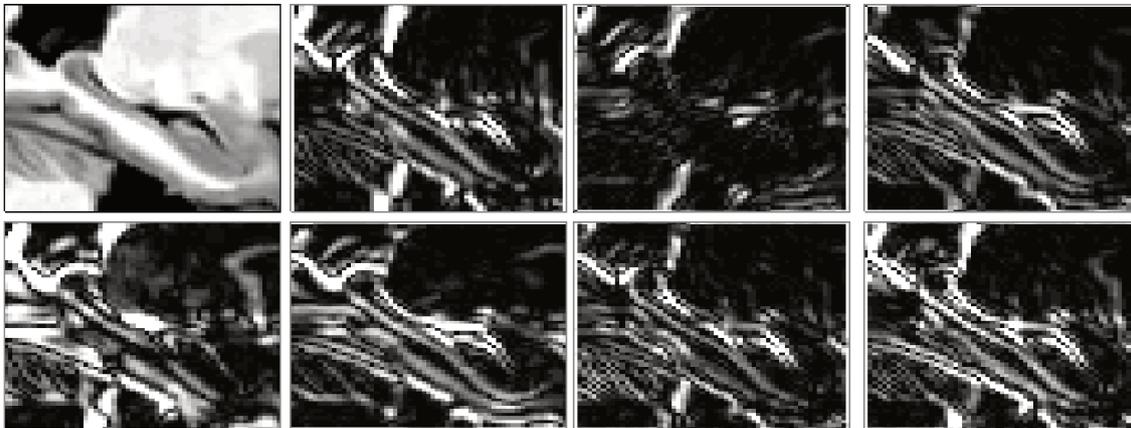


Figure 5.2: Sample of absolute value of complex number computed by Gabor filters. The hippocampal area is filtered with gabor at different directions. From left: Original image, textural maps provided by our Gabor filters bank given by the parameter set $\{(\theta = 0; \phi = 0), (\pi/2; \pi/2), (0; \pi/2), (0; \pi/4), (0; -\pi/4), (-\pi/4, \pi/4), (\pi/4, \pi/4)\}$. It can be noticed that for each direction, different textures are extracted.

5.3.3 Weak classifiers aggregation

In previous works on patch-based grading, the weak classifier aggregation was performed using a simple averaging (Coupé et al., 2012b; Tong et al., 2017a). While using a strategy based on averaging enables to be robust to noise, this may remove relevant information on weak classifiers distribution. Therefore, in this paper we propose to approximate weak classifiers distribution using histogram. Consequently, we classify histogram bins instead of classifying mean grading value over the segmentation mask. Here, histograms were separately estimated for right and left hippocampus and concatenate into a single feature vector. The number of bins are set following Sturge’s formula that intends to find the optimum number of bin

for a certain size of sample (Sturges, 1926). Finally, to prevent bias introduced by structure alterations related to aging, all the grading values are age corrected with a linear regression based on the CN group (Dukart et al., 2011). This correction is done by removing the test CN subjects into a cross-validation procedure.

5.3.4 Implementation

During our experiments, texture maps were obtained using one scale and 11 different directions. The texture-based grading maps were estimated using patches of $5 \times 5 \times 5$ voxels. The grading step based on the optimized patch match was performed using $K = 50$. The required computational time was 3s per texture maps, thus the global grading step required 15 seconds with our setup. A support vector machine (SVM) with a linear kernel was used to classify each test subject. We used the Matlab function provided by the Statistics and Machine Learning Toolbox. In our experiments, the soft margin parameter C was optimized with a Bayesian optimization method. The results of each experiment were compared in terms of accuracy (ACC) and area under the ROC curve (AUC), specificity (SPE), and sensitivity (SEN). The AUC is estimated with the *a posteriori* probabilities provided by the SVM classifier. We carried out several experiments: CN versus AD, CN versus pMCI, AD versus sMCI and sMCI versus pMCI. A t-test were performed to study the significance of the results provided by adaptive fusion scheme compared to mean of textural maps and late fusion into SVM classifier. Finally, our new texture-based grading framework was validated within a repeated stratified 10-fold cross-validation procedure iterated 50 times for CN versus AD, CN versus pMCI and AD versus sMCI comparisons. The mean ACC, AUC, SPE, and SEN over these 50 iterations are provided as results. As demonstrated in (Tong et al., 2017a), training the classifier with CN and AD enables to discriminate sMCI and pMCI subjects better. Moreover, it enables to perform classification without cross-validation procedure and to limit bias and over-fitting problem. Therefore, only one run was performed for sMCI versus pMCI comparison.

5.4 Results

5.4.1 Optimal number of directions

First, the optimal number of filter directions were investigated. Figure 5.3 shows the evolution of accuracy related to the number of directions. This experiment demonstrates that 5 different directions are enough to obtain the best results for CN versus AD comparison. Indeed, the accuracy does not increase using more directions. The best accuracy is reached with 5 different directions for sMCI versus pMCI. A fusion of Gabor filters at different scales was also performed. However, this experiment shown that filters at the full image resolution is enough to obtain

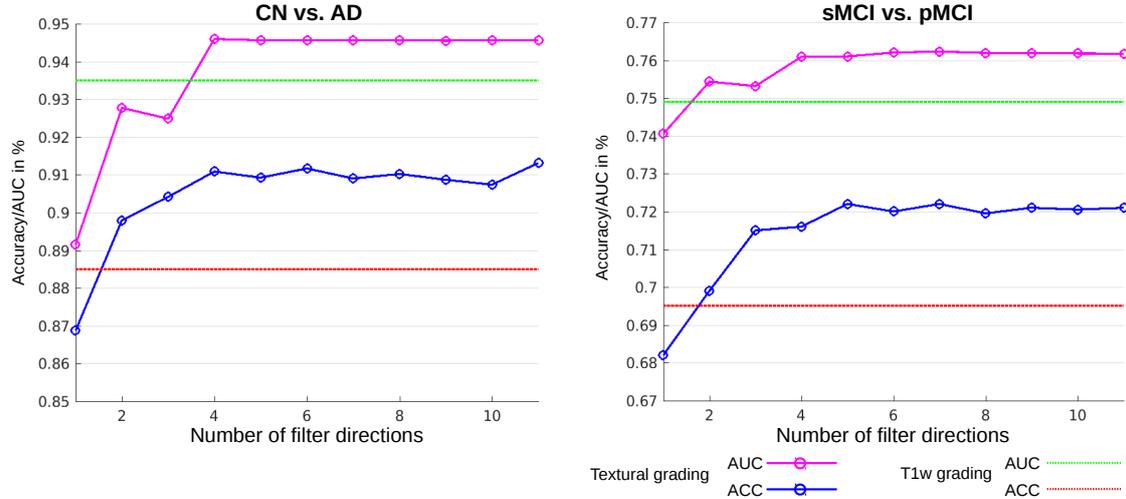


Figure 5.3: Evolution of the accuracy and AUC for CN versus AD and sMCI versus pMCI comparisons according to the number of Gabor filters. Red and green dotted line represented the accuracy and the AUC of intensity-based grading.

the best results. The experiments showed that the optimal set of filters directions is $\{(\theta = 0; \phi = 0), (\pi/2; \pi/2), (0; \pi/2), (\pi/4; \pi/4), (\pi/4; -\pi/4)\}$. Therefore, in the rest of the experiments, comparisons were performed with Gabor filters in these 5 different directions and at the full images resolution.

5.4.2 Comparison grading based on intensity vs. texture

To estimate the improvement provided by texture-based approach, we compare results obtained with our framework using intensities of the images in the MNI space (*i.e.*, intensity-based grading) and texture maps. For this comparisons, intensity and texture-based grading were estimated using exactly the same pipeline involving

Table 5.2: Comparison of grading features based on histogram representation of the probability distribution of grading values. This table presents a comparison of intensity-based grading and texture-based grading. These results show that texture-based grading improves AUC of all comparisons. Moreover, these results show that histogram representation provides similar or better results for all comparisons than using average value (see Table 5.3). All the results are expressed in percentage of AUC, SEN, and SPE.

	Intensity-based grading histo. (AUC/SEN/SPE in %)	Texture-based grading histo. (AUC/SEN/SPE in %)
CN vs. AD	93.5/ 95.5 /82.7	94.6 /94.2/ 86.6
CN vs. pMCI	90.0/81.8/ 81.4	92.0 / 92.5 /81.2
AD vs. sMCI	81.1/ 78.5 /68.3	82.6 /77.6/ 72.6
sMCI vs. pMCI	74.9/ 77.6 /67.2	76.1 /74.9/ 70.2

Table 5.3: Comparison of grading features based on mean of the grading values within the hippocampus structure. This table presents a comparison of intensity-based grading and texture-based grading. These results show that texture-based grading improves AUC of all comparisons. All the results are expressed in percentage of AUC, SEN, and SPE.

	Intensity-based grading mean (AUC/SEN/SPE in %)	Texture-based grading mean (AUC/SEN/SPE in %)
CN vs. AD	92.6/86.7/83.3	94.7/93.4/87.6
CN vs. pMCI	89.9/78.2/ 85.4	92.3/91.6/83.0
AD vs. sMCI	80.8/76.2/69.9	82.2/77.6/71.0
sMCI vs. pMCI	73.2/76.4/ 65.0	75.1/77.0/64.1

adaptive fusion and histogram-based weak classifiers aggregation. Table 5.2 summarizes the results of intensity-based grading and the proposed texture-based grading obtained with 5 Gabor filters. Results are expressed with area under the curve (AUC), sensibility (SEN) and specificity (SPE) measures.

As it is shown, texture-based grading improves classification performances in all experiments using mean or histogram-based grading. Indeed, the comparisons conducted with histogram-based representation show that texture-based grading obtains 94.6% of AUC for CN versus AD, 92.0% of AUC for CN versus pMCI, and 82.6% of AUC for AD versus sMCI comparisons while intensity-based grading obtains 93.5% of AUC for CN versus AD, 90.0% of AUC for CN versus pMCI, and 81.1% of AUC for AD versus sMCI comparisons. Finally, with histogram representation, texture-based grading obtains 76.1% of AUC for sMCI versus pMCI comparisons and intensity-based grading obtains 74.9%. As results based on histogram representation, the average grading aggregation follows the same tendency. These results demonstrate that texture maps enable to better capture structural alterations.

5.4.3 Comparison average grading vs. histogram-based grading

In this section, we compare our proposed histogram-based weak-classifier aggregation of texture-based grading values with a straightforward average that is usually used in patch-based grading framework. As presented in Table 5.3 and 5.2, the results show that histogram representation of weak classifiers distribution provides similar or better classifications results for all comparisons. Histogram-based aggregation obtains 94.6% and 92.0% of AUC while the average obtains 94.7% and 92.3% of AUC for CN versus AD and CN versus pMCI comparisons, respectively. Moreover, histogram-based aggregation obtains better results for AD versus sMCI and sMCI versus pMCI comparisons with 82.6%, and 76.1% of AUC compare to the average that obtains 82.2% and 75.1% of AUC for the same comparisons, respectively.

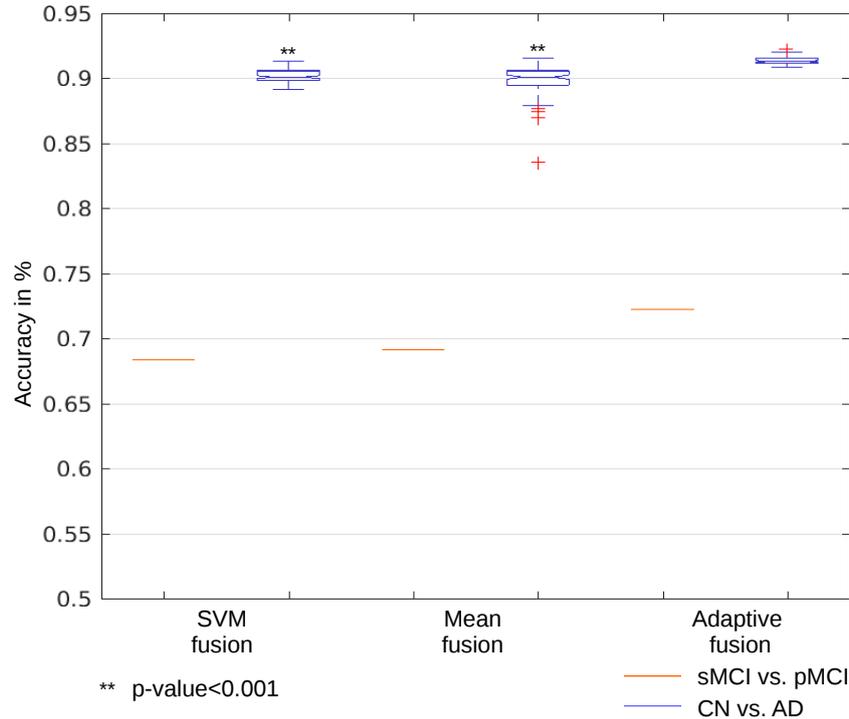


Figure 5.4: Comparison of different fusion schemes: mean of textural-grading maps, adaptive fusion of textural-grading maps and fusion into a SVM classifier (*i.e.*, concatenation of the histogram representing the grading of each texture direction as the input of SVM method). This comparison shows that adaptive fusion provides best results for both AD detection and prediction (*i.e.*, CN versus AD and sMCI versus pMCI). P-values were estimated with a t-test to compare adaptive fusion with other fusion methods. ** indicates that the p value is inferior than 0.001.

5.4.4 Comparison of different fusion schemes

Our fusion scheme was compared with a fusion based on the mean of texture-based grading maps (*i.e.*, fusion of the different grading maps provided at each direction with a straightforward average) and a SVM fusion of our texture-based grading features (*i.e.*, concatenation of the histogram features at the different considered directions into the SVM classifier). Results are summarized in Figure 5.4. During the experiments, adaptive fusion obtained an accuracy of 91.3% for CN versus AD comparison, the fusion using SVM classifier obtained 90.1% and the mean fusion obtained 89.1%. Moreover, for sMCI versus pMCI comparison adaptive fusion obtains 72.2% of accuracy while SVM fusion obtains 68.3% and mean fusion obtains 69.1%. Thus, adaptive fusion obtained the best results. Indeed, the results obtained by adaptive fusion is 1.2 percentage point higher than SVM fusion and 2.1 percentage point higher than mean fusion for CN versus AD comparison and 3.8 percentage point higher than SVM fusion and 3.1 percentage point higher than mean fusion. In order to study the significance of the accuracy differences between each fusion

method, p-values were estimated with a t-test. These experiments showed that adaptive fusion provides significantly better results for CN versus AD comparisons.

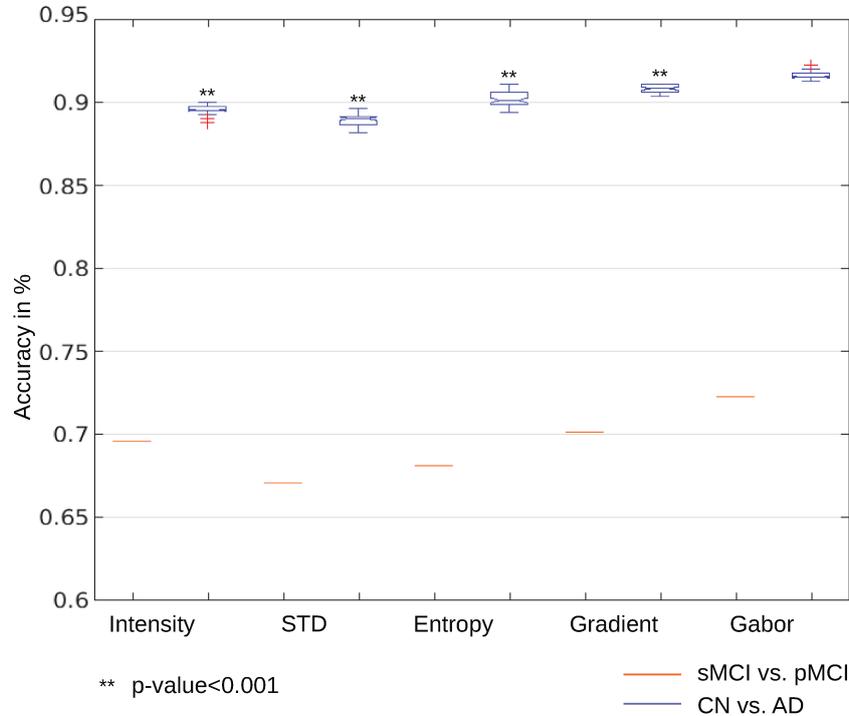


Figure 5.5: Comparison of classification performances for different texture filters. From left to right, we compare standard deviation (STD), entropy, gradient and Gabor filters. Entropy and STD filters were computed with a window size of $5 \times 5 \times 5$ voxels. The results shows that only gradient and Gabor filters enable to improve performances for both AD detection and prediction. P-values were estimated with a t-test to compare Gabor filters to other texture filters. ** indicates that the p value is inferior than 0.001.

5.4.5 Comparison of different texture-based filters

Figure 5.5 shows comparison of classification results for different texture filters. This comparison was conducted with standard deviation (STD), entropy, gradient, and Gabor filters. Entropy and STD filters were computed into a window size of $5 \times 5 \times 5$ voxels. Patch-based grading maps from gradient and Gabor filters were merged with our novel adaptive fusion scheme. STD filters obtained 89.1% of accuracy for CN versus AD comparison and 67.1% for sMCI versus pMCI comparison. Entropy filter obtained an accuracy of 90.5% for CN versus AD comparison and 68.1% for sMCI versus pMCI comparison. Gradient filter obtained 90.8% for CN versus AD and 70.1% for sMCI versus pMCI. This experiment shows that only gradient and Gabor filters obtain better results for both, AD detection and prediction compared to intensity-based grading with 91.3% and 72.2% of ACC for CN versus AD and

sMCI versus pMCI comparisons, respectively. Thus, patch-based grading applied on an optimal set of Gabor filters provides better results than others texture filters for both considered comparison. Gabor filters improve by 3.2 and 2.1 percent points of accuracy for CN versus AD and sMCI versus pMCI, respectively, compared to gradient filter.

5.4.6 Comparison to state-of-the-art methods

In addition, a comparison with state-of-the-art methods are provided in Table 5.4. The results of this comparison are expressed in accuracy. On one hand, to compare classification results using the same structure, the proposed framework is compared with grading methods based on hippocampus (see the upper part of Table 5.4). Thus, our proposed texture-based grading method is compared with the original patch-based grading method (Coupé et al., 2012b), a grading based on multiple instance learning method (Tong et al., 2014), and a patch-based grading based on a sparse representation using two different registration strategies (Tong et al., 2017a). This comparison shows that our method provides best results among hippocampus-based grading methods. It reaches 91.3% of accuracy for CN versus AD, and 72.2% of accuracy for sMCI versus pMCI comparisons. On the other hand, our proposed method applied into hippocampus is compared with methods based on a whole brain analysis using similar dataset (see the lower part of Table 5.4). Indeed, we compare our texture-based grading approach applied on hippocampus with a patch-based grading method based on a sparse representation applied on the whole brain (Tong et al., 2017a), a sparse ensemble grading method that analyzes the whole brain (Liu et al., 2012), and a Deep Learning (DL) method based a whole brain analysis (Suk et al., 2017). The results show that our method obtains the best accuracy for AD versus CN. This result is similar to classification results obtained with a DL and sparse ensemble grading method (Suk et al., 2017; Liu et al., 2012). However, methods based on a whole brain analysis and using non linear registration obtain more accurate classification results for sMCI versus pMCI.

5.5 Discussion

In this work, to improve patch-based grading framework, we proposed to capture texture information with a bank of Gabor filters. Our experiments showed that using more than 5 directions does not improve the results while increasing computational time (see Figure 5.3). Moreover, we also investigated a multi-scale texture approach. However, the experiments carried out showed that only one scale, at the full image resolution, is enough and the use of multi-scale texture did not improve classification performances. Therefore, we propose a multi-directional texture-based grading framework based on 1 scale and 5 directions.

A new grading values aggregation method based on histogram was also proposed.

Table 5.4: Comparison with state-of-the-art methods, all the results are expressed in accuracy. The upper part of this table presents results of results of methods applied on the hippocampus structure. Texture-based grading improves CN versus AD classification by 2.3 percent points and sMCI versus pMCI classification by 1.2 percent points of accuracy compared to grading approaches based on the same hippocampus structure. The lower part presents results of methods that propose a whole brain analysis. Compared to these approaches, our method obtains similar results than advanced methods based on whole brain analysis for CN versus AD classification. However, such methods obtain better performances for sMCI versus pMCI classification. In this table, we provide the type of registration (Reg.) involved in the methods and the type of features (Feat.).

Methods	Registration	Feature	CN vs. AD (ACC in %)	sMCI vs. pMCI (ACC in %)
Hippocampus				
(Coupé et al., 2012b)	Affine	Intensity	88.0	71.0
(Tong et al., 2014)	Affine	Intensity	89.0	70.0
(Tong et al., 2017a)	Affine	Intensity	–	66.0
(Tong et al., 2017a)	NL	Intensity	–	69.0
Proposed Method	Affine	Texture	<u>91.3</u>	72.2
Whole Brain				
(Tong et al., 2017a)	Affine	Intensity	–	66.7
(Tong et al., 2017a)	NL	Intensity	–	75.0
(Liu et al., 2012)	NL	GM	90.8	–
(Suk et al., 2017)	NL	GM	91.0	74.8

GM = Grey matter

NL = Non linear

During our experiments, histogram representation of grading values distribution did not provide improvement for CN versus AD comparison compared to use a simple average value. That could be explained by the fact that CN and AD distributions are well separated and a parametric representation of their distributions is enough to discriminate these two groups. However, for sMCI versus pMCI case, the two distributions are less separable and histogram representation lead to better classification performances with in average a gain of 1 percentage point of AUC compared to a simple average value.

In order to fuse efficiently the different texture-based grading maps, we proposed a novel patch-based grading fusion scheme. This method is based on a confidence value estimated at each voxel. The comparison with different fusion schemes demonstrated the efficiency of our method. Indeed, compared to a straight average of the texture-based grading maps and a SVM fusion of the final histogram features with the classifier, our proposed method obtained best accuracy for AD detection and prediction. Moreover, the obtained improvement was significant. This improvement

can be explained by the fact that our proposed adaptive fusion method weights the grading values of each texture map according to their relevances while the fusion into SVM classifier and the average of texture maps considers each grading value as having the same importance.

Our work hypothesis is also that directional texture filters enable to improve patch comparison, and thus increase AD detection and prediction accuracy. To validate this hypothesis, our novel texture-based grading using an optimal set of Gabor filters were compared with others texture filters as done for segmentation in (Wachinger et al., 2017) (see Figure 5.5). STD, entropy, gradient and Gabor filters were compared for AD detection and prediction. This experiment showed that STD and entropy does not enable to improve patch comparison compare to intensity. The limitation of these filters might be to perform feature estimation within a window. Thus, only gradient and Gabor filters improve classification performances for AD detection and prediction. Moreover, Gabor filters obtain best results for both comparisons. This improvement is related to the use of additional texture directions compared to the three texture directions provided by gradient filter.

Table 5.4 summarizes the comparison of our proposed method with other grading methods proposed in the literature. These results demonstrate that Gabor filters enable to better capture structural alterations than method based on intensity or grey matter data. Indeed, texture maps provide enhance information leading to a better grading process. Thus, our method outperforms other grading methods using intensity when applied on the same structure (Coupé et al., 2012b; Tong et al., 2014, 2017a). At the lower part of Table 5.4, we compare the performance of our hippocampus-based grading method with methods using the whole brain. First, for AD versus CN, the proposed method obtained similar or better results than methods applied over the whole brain. It is important to note that these methods require non linear registration (Liu et al., 2012; Tong et al., 2017a; Suk et al., 2017) while our method only requires affine registration and proposes a fast grading step. Second, for sMCI versus pMCI, our method obtained better results than all the methods involving a simple affine registration, including whole brain method proposed in (Tong et al., 2017a). On the other hand, the best results for sMCI versus pMCI are produced by whole brain grading (Tong et al., 2017a) using non linear registration. The improvement when using non linear registration is observed for hippocampus-based and whole brain methods (Tong et al., 2017a). However, this improvement is obtained at the expense of using non linear registration, which is subject to failure and requires high computational time. Our method also demonstrated competitive performances for AD versus CN classification compared to advanced DL methods using whole brain and non linear registration (Suk et al., 2017). Finally, this comparison shows that patch-based grading methods (Tong et al., 2017a) obtain similar or better results than recent deep learning methods (Suk et al., 2017) when applied over the entire brain after non linear registration.

5.6 Conclusion

In this chapter, we have proposed a new texture-based grading framework to better capture structural alterations caused by AD. Our method combines textural grading maps estimated on texture maps with our new adaptive fusion scheme. Moreover, we also have proposed an histogram-based weak classifiers aggregation approach to better discriminate early stages of AD. We have studied the optimal set of texture directions. Experiments conducted in this work demonstrated the relevance of using textural information in combination with with our novel locally adaptive fusion method. Finally, we have demonstrated the competitive performances of our new texture-based grading framework compared to several state-of-the-art approaches.

Chapter 6

Multimodal-based grading

6.1	Introduction	78
6.2	Materials	79
6.3	Experiments	79
6.3.1	MRI processing	79
6.3.2	DTI processing	80
6.3.3	Features estimation	81
6.3.4	Implementation	81
6.3.5	Validation	81
6.4	Results	82
6.4.1	Modality comparisons	83
6.4.2	Relationship with cognitive scores	84
6.4.3	Comparison with state-of-the-art methods	84
6.5	Discussion	86
6.6	Conclusion	86

6.1 Introduction

To date, patch-based grading has been applied on structural MRI, using intensity feature (Coupé et al., 2012b; Tong et al., 2017a), grey matter probability maps (Liu et al., 2012; Komlagan et al., 2014), and, as we proposed in the previous chapter, using texture filters (see chapter 5).

However, although structural MRI is a valuable imaging technique to measure global structural modifications, such modality is not able to capture microstructural degradation. Moreover, the microstructural modifications caused by AD are considered to occur before the atrophy measured by structural MRI. Therefore, diffusion MRI (d-MRI) appears as a potential candidate to detect the earliest sign of AD. Several diffusion tensor imaging (DTI) studies proposed automatic methods to detect modifications of diffusion parameters into the whole white matter using machine learning (O'Dwyer et al., 2012; Dyrba et al., 2013, 2015). Others studies showed modifications of diffusion parameters for AD patients into specific white matter structures such as corpus callosum (Nir et al., 2013; Wang et al., 2015), fornix (Liu et al., 2011), cingulum (Nir et al., 2013) and also in grey matter tissue such as hippocampus (Rose et al., 2008). More advanced d-MRI studies using brain connectivity and fiber tracking have been proposed to extract features describing axonal fibers alterations (Liu et al., 2011; Wee et al., 2012; Prasad et al., 2015). Finally, it has been shown that hippocampal mean diffusivity (MD) is correlated to pathology progression and thus could be used as an efficient biomarker of AD (Fellgiebel and Yakushev, 2011). Moreover, it was demonstrated that MD increases with the development of AD in the grey matter (Kantarci et al., 2005; Müller et al., 2005; Fellgiebel et al., 2006). Some methods proposed to fuse d-MRI and s-MRI biomarkers to use the complementarity of these two MRI modalities (Cui et al., 2012; Li et al., 2014). These studies showed the complementarity of s-MRI and d-MRI to capture early alteration led by AD.

Therefore, in this chapter, we propose to extent our patch-based framework to DTI grading and multimodal grading. First, we will show that patch-based features applied on DTI demonstrates competitive performances to classify the early stages of AD. Second, we propose to study the alterations of hippocampus with a multimodal patch-based grading framework based on the adaptive fusion scheme described in the Chapter 4. The conducted experiments show that patch-based grading method based on d-MRI enables to better capture alterations at the first stage of AD but does not improve results for late detection of AD. However, the fusion of s-MRI and d-MRI provides hippocampus biomarker obtaining best performances for AD detection and prediction.

6.2 Materials

Data used in this work were obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) dataset¹. This dataset includes AD patients, MCI and control normal (CN) subjects. The group of MCI is composed of subjects who have abnormal memory dysfunctions. In this work we used data from the ADNI2 campaign that proposes eMCI and IMCI stages. The eMCI and IMCI subgroups were obtained with the Wechsler Scale-Revised Logistical Memory I and II tests in accordance with the education levels of each subject.² ADNI2 provides T1-weighted (T1w) MRI and DTI scans for 54 CN, 79 eMCI, 39 IMCI and 47 AD subjects. Only patients whose have T1w and DTI were selected in our work. Hence, in this work we used 52 CN, 99 MCI composed of 65 eMCI, 34 IMCI and 38 AD instead of the whole initial ADNI2 dataset. All MRI data and clinical status were collected at the baseline. The list of subjects involved in our experiments is available online³. Table 6.1 shows the distribution of the data for each group. The s-MRI and d-MRI scans used for all considered subjects in this study were acquired with the same protocol⁴. T1w MRI acquisition protocol had been done with the 3D accelerated sagittal IR-SPGR, according to the ADNI protocol (Jack et al., 2008). The d-MRI were composed of 46 separate angles, 5 T2-weighted images with no diffusion sensitization (b0 images) and 41 directions (b=1000s/mm²). The d-MRI protocol was chosen to optimize the signal-to-noise ratio in a fixed scan time (Jahanshad et al., 2010). The native resolution of s-MRI and d-MRI was set to 1mm³ and 2mm³, respectively.

6.3 Experiments

6.3.1 MRI processing

As described in the chapter 5, T1w images were processed using the volBrain system (Manjón and Coupé, 2016)⁵. This system is based on an advanced pipeline providing automatic segmentation of different brain structures from T1w MRI. The preprocessing is based on: (a) a denoising step with an adaptive non-local mean filter (Manjón et al., 2010), (b) an affine registration in the MNI space (Avants et al., 2011), (c) a correction of the image inhomogeneities (Tustison et al., 2010) and (d) an intensity normalization.

¹<http://adni.loni.ucla.edu>

²<http://adni.loni.usc.edu/wp-content/uploads/2008/07/adni2-procedures-manual.pdf>

³http://www.labri.fr/~khett/dataset_scirep18.csv

⁴https://adni.loni.usc.edu/wp-content/uploads/2010/05/ADNI2_GE_3T_22.0_T2.pdf

⁵<http://volbrain.upv.es>

Table 6.1: Description of the dataset used in this work. Data are provided by ADNI. MMSE:Mini-Mental State Examination; CDR-SB: Clinical Dementia Rating-Sum of Boxes; RAVLT:Rey’s Auditory Verbal Learning Test; FAQ: Functional Activity Questionnaire; ADAS(11/13): Alzheimer’s Disease Assessment Scale.

	CN	eMCI	IMCI	AD	P value
Number	52	65	34	38	
Age (years)	72.6 \pm 5.9	73.0 \pm 7.7	73.5 \pm 6.6	73.8 \pm 8.7	p = 0.80 ^a
Gender (F/M)	29/23	39/26	21/13	20/18	$\chi^2=3.12$, p = 0.37 ^b
MMSE	28.9 \pm 1.2	28.2 \pm 1.5	27.3 \pm 1.8	23.4 \pm 1.7	p < 0.01 ^{a*}
CDR-SB	0.0 \pm 0.1	1.2 \pm 0.6	1.7 \pm 0.8	4.6 \pm 1.4	p < 0.01 ^{a*}
RAVLT	45.4 \pm 9.7	36.5 \pm 10.2	30.7 \pm 8.9	22.6 \pm 7.0	p < 0.01 ^{a*}
FAQ	0.2 \pm 0.9	2.3 \pm 3.7	4.3 \pm 4.8	14.6 \pm 6.6	p < 0.01 ^{a*}
ADAS11	5.2 \pm 3.0	8.1 \pm 3.6	12.5 \pm 4.9	20.2 \pm 7.6	p < 0.01 ^{a*}
ADAS13	8.4 \pm 4.4	13.3 \pm 5.4	20.2 \pm 6.7	30.0 \pm 9.0	p < 0.01 ^{a*}

* Significant at p < 0.05.

^a Chi-square test (df = 3).

^b Kruskal–Wallis test (df = 3).

6.3.2 DTI processing

The preprocessing of the diffusion weighted images is based on: (a) a denoising step based on the LPCA filter (Manjón et al., 2013) and (b) a correction of the head motion using an affine registration. Afterwards, we performed several steps to first obtain the mapping between the DWI native space and the MNI space and then to estimate the MD in the MNI space.

- 1) Estimation of the mapping between DWI native space and MNI space: First, a diffusion tensor model (Basser et al., 1994) estimated at each voxel using Dipy library (Garyfallidis et al., 2014). The obtained MD is first linearly registered to the CSF map obtained from the T1w in the MNI space. Then, the MD (in the MNI space) is non-linearly registered to the CSF map (in the MNI space) to compensate for echo-planar imaging (EPI) distortions (Avants et al., 2011). Afterwards, the affine transformation and the non-linear deformations are concatenated into a single transformation to obtain the final mapping (including EPI distortion correction) from the DWI native space to the MNI space. It has to be noted that the MD map estimated in the DWI native space is only used to estimate the mapping between both spaces.
- 2) Estimation of the MD in the MNI space: First, the deformation field estimated at the previous step is used to registered the b_0 and each DWI direction from their native space into the MNI space using b-spline interpolations (Avants et al., 2011). This is done to limit interpolation artifacts and to correct partial volume effect (PVE). Indeed, it has been shown that up-sampling each DWI

direction individually using interpolation before estimating DTI parameters enables to greatly reduce PVE present in DTI (Dyrby et al., 2014). Thus, the final diffusion tensor model is estimated in the MNI space using all the non-linearly registered DWI and b_0 .

To analyze microstructural modifications, the MD is estimated within the hippocampus structure with the segmentation described in the previous section. MD is defined as $\frac{\lambda_1 + \lambda_2 + \lambda_3}{3}$ where $\lambda_1, \lambda_2, \lambda_3$ are the three eigenvalues of the fitted tensor.

Finally, a quality control was conducted to exclude data presenting miss-segmentation or miss-registration after MRI and DTI preprocessing step. Thus, 10 CN subjects, 18 eMCI, 5 IMCI, and 9 AD patients has been excluded from the initial considered ADNI2 dataset (see the dataset used in our experiments Table 6.1).

6.3.3 Features estimation

Features were estimated over the right and left hippocampus masks. To reduce the inter-individual variability, all volumes are normalized by the total intra-cranial volume (Whitwell et al., 2001). Afterwards, we aggregate local weak classifiers of the grading map into a single feature for both hippocampus (*i.e.*, right and left) by averaging them. Therefore, patch-based grading features are computed by an unweighted vote of the weak classifiers using the segmentation masks (see Figure 6.1). Finally, to prevent the bias introduced as the structure alterations due to aging, all the features (*i.e.*, volume, mean of MD and MPBG) are age corrected with a linear regression based on the CN group (Dukart et al., 2011).

6.3.4 Implementation

To find the most similar patches in the training library, we use the OPAL method (Giraud et al., 2016). OPAL is a fast approximate nearest neighbor patch search technique. This method enables to process each modality in about 4 seconds on a standard computer. The training library is equally composed of 37 images for both CN and AD subjects, leading to $|T| = 76$. The number of patches extracted from each the training library is $K = 160$ (*i.e.*, 80 from CN subjects and 80 from AD patients) and the patch size is $5 \times 5 \times 5$ voxels. Furthermore, we used zero normalized sum of squared differences for T1w to compute the L2 norm (see Equation (3.2)). On the other hand, d-MRI is a quantitative imaging technique. Therefore, to preserve the quantitative information, a straight sum of squared differences is used for MD in Equation (3.2),

6.3.5 Validation

To evaluate the efficiency of each considered biomarker to detect AD alterations, CN group is compared to AD patients group. In addition, to discriminate the impairment severity of MCI group, eMCI versus IMCI classification is conducted.

The classification step is performed with a linear discriminant analysis (LDA) within a repeated stratified 5-fold cross-validation iterated 200 times. Mean area under the curve (AUC) and mean accuracy (ACC) are computed to compare performance for each biomarker over the 200 iterations.

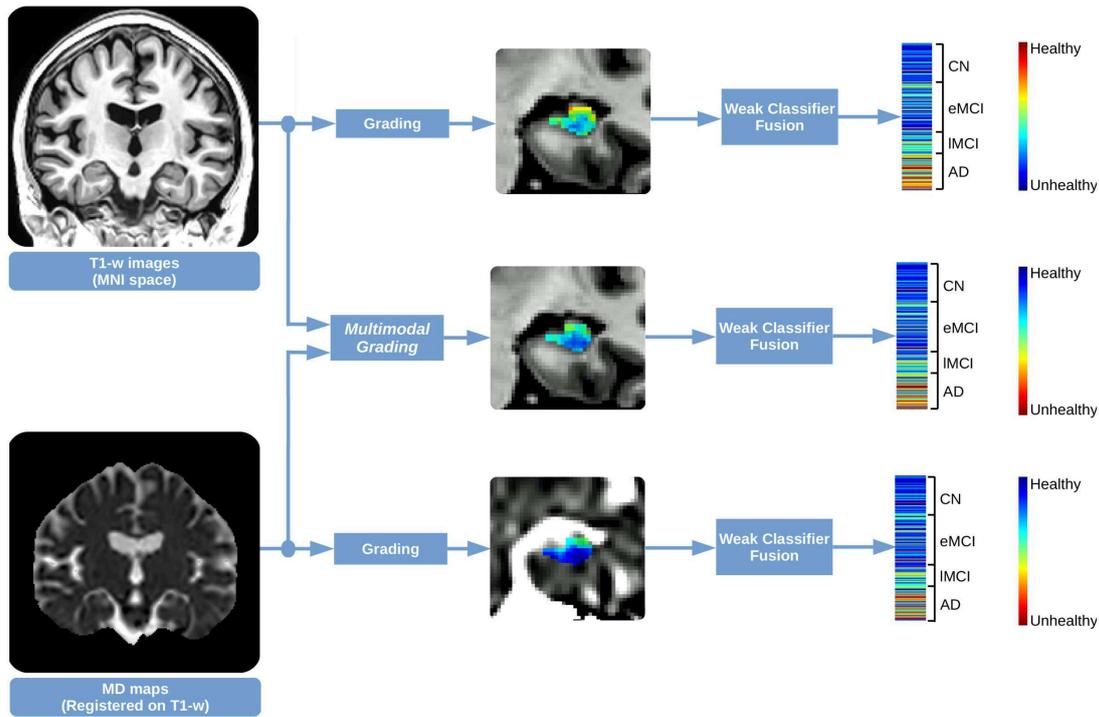


Figure 6.1: Proposed multimodal patch-based grading framework. At left, the input data: T1w images registered into the MNI space and MD maps registered on the T1w images. At the middle: the corresponding coronal view of a PBG and MPBG maps estimated on T1w and MD. At right, the considered hippocampal biomarkers for all subjects under study.

6.4 Results

In this section, the results are presented in three parts. In the first part, we compare the different approach applied within the entire hippocampus structure to evaluate the performance of our new MPBG compared to usual biomarkers such as volume and average MD. Afterward, we compare the results of our proposed multimodal biomarker with state-of-the-art methods based on d-MRI to show the competitive performance of our approach. Finally, in a last part, we study the relationship of our multimodal biomarker with different cognitive tests.

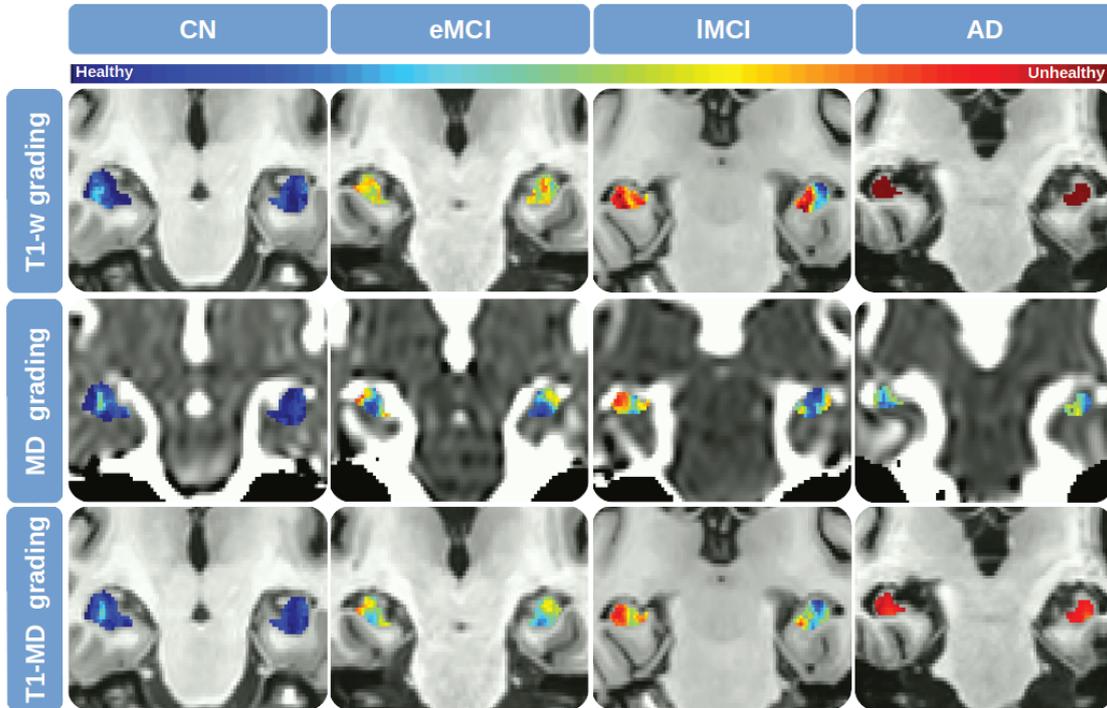


Figure 6.2: Results obtained for different severities of cognitive impairments. From top to bottom slices on the coronal plane of the segmentation maps and the fusion of T1w and MD patch-based grading with the proposed multimodal patch-based grading method. The blue and the red colors represent the healthy and altered tissues, respectively.

Table 6.2: Mean AUC of the different features estimated over the whole hippocampal structure. In bold font the best result for each specific comparison. All results are expressed in percent.

Method	CN vs. AD	eMCI vs. IMCI
Volume	86.6	59.4
MD	80.6	55.6
T1w PBG	92.6	67.5
MD PBG	89.2	69.5
MPBG	92.1	69.5

6.4.1 Modality comparisons

Results of the comparisons over the hippocampus are represented in Table 6.2. In this experiment, we compared the results of volume, mean of MD and PBG applied with both modality and MPBG.

First, the hippocampus volume and its average of MD were compared. For CN versus AD classification the volume obtains 86.6% of AUC and the average of MD obtains 80.6%. For eMCI versus IMCI classification the volume and the average of

MD obtain 59.4% and 55.6% of AUC, respectively. Experiments demonstrate that the hippocampus the volume obtains better classification results than the average of MD for all comparison, especially for CN versus AD. Second, PBG biomarkers applied with T1w and MD were compared. The results showed that T1w PBG provides better results than MD PBG with 92.6% of AUC for CN versus AD classification. However, for eMCI versus IMCI classification MD grading provides the best results with 69.5% of AUC. MPBG methods combining both modalities reaches the best results for CN versus AD and eMCI versus IMCI with 92.1% and 69.5% of AUC, respectively. Finally, the proposed MPBG biomarker provides results similar to the best modalities for all considered comparisons. Compared to volume, MPBG improves CN versus AD comparison result by 5.5% of AUC and by over 10% of AUC for eMCI versus IMCI comparison. Thus, MBPG biomarker has a good capability to capture modifications caused by AD at different severity stages (see Figure 7.3).

6.4.2 Relationship with cognitive scores

To investigate relationships between cognitive scores and MPBG values, we performed a generalized linear analysis with the following model: $MPBG = \beta_0 + \beta_1.age + \beta_2.sex + \beta_3.MMSE + \beta_4.RAVLT + \beta_5.FAQ + \beta_6.CDRSB + \beta_7.ADAS11 + \beta_8.ADAS13$. We found significant relationship of hippocampal MPBG with sex ($p < 0.01$), MMSE ($p < 0.05$) and ADAS 13 ($p < 0.01$).

Table 6.3: This table presents the p-values of our multimodal grading biomarker with some meta-data (*i.e.*, sex, age, and cognitive tests). * and † indicate a p value inferior than 0.05 and 0.01, respectively.

	Sexe	Age	MMSE	RAVLT	FAQ	CDRSB	ADAS11	ADAS13
Hippocampus	0.002 †	0.638	0.030 *	0.134	0.223	0.135	0.050 *	0.007 †

6.4.3 Comparison with state-of-the-art methods

To evaluate the performance of the proposed MPBG, we compared it with state-of-the-art multimodal methods using d-MRI. To this end, we used the ACC values published by the authors. Table 6.4 shows the comparison of our proposed biomarker within the hippocampus providing the best results with state-of-the-arts methods using similar dataset based on ADNI2. We compared these biomarkers with a method using features based on tractography (Nir et al., 2015), a method based on a connectivity network of the different brain structures (Prasad et al., 2015), and a voxel-based method that analyzes alterations of white matter (Maggipinto et al., 2017). The results of comparison show that MPBG over whole hippocampus obtains the best score for AD versus CN with 88.1% of accuracy while the best result is achieved by a voxel-based method with a feature selection (Maggipinto et al., 2017) that obtained 87.0% on similar ADNI2 dataset. To the best of our knowledge, the

Table 6.4: Comparison of our proposed MPBG biomarkers with state-of-the-arts methods based on s-MRI and d-MRI using a similar ADNI2 dataset. All results are expressed in percentage of accuracy.

Method	Subjects				Feature	Classifier	Classification ACC	
	CN	eMCI	IMCI	AD			CN/AD	eMCI/IMCI
(Nir et al., 2015)	44	74	39	23	Tractography	SVM	84.9%	n/a
(Prasad et al., 2015)	50	74	38	38	Connectivity network	SVM	78.2%	63.4%
(Zhan et al., 2015)	n/a	73	39	n/a	Connectivity network	SLG	n/a	65.0%
(Maggipinto et al., 2017)	50	22	18	50	Voxel-based	RF	87.0%	n/a
(La Rocca et al., 2018)	52	85	38	47	Connectivity network	RF	83.0%	n/a
MPBG hippocampus	62	65	34	38	Patch-based	LDA	88.1%	68.8%

LDA = Linear Discriminant Analysis

SLG = Sparse Logistic Regression

SVM = Support Vector Machine

RF = Random Forest

only works providing eMCI and lMCI comparison using d-MRI from similar ADNI2 dataset are based on a connectivity network and obtained 63.4% and 65.0% (Prasad et al., 2015; Zhan et al., 2015). The difference between the two approaches mainly comes from the nature of the final considered features and the classification step. Indeed Prasad et al. (2015) uses network measures proposed by Rubinov and Sporns (2010) as input of a linear SVM method, while Zhan et al. (2015) based the decision on the combination of a single value decomposition and a sparse logistic regression (SLR) using the raw connectivity network features. These comparisons demonstrate the relevance of MPBG biomarkers for AD detection and prediction. Indeed, our method provides similar results than the best methods with similar dataset for CN versus AD classification and provides the best results for eMCI versus lMCI classification. Moreover, the proposed MPBG within the hippocampus obtains competitive performances for eMCI versus lMCI classification with an accuracy of 68.8%, that increases by almost 4 points of percentage compared to connectivity network-based methods.

6.5 Discussion

The major contribution of the work presented in this chapter is a multimodal patch-based grading method applied to the hippocampus for the early detection of AD. Since PBG method applied into s-MRI and d-MRI have obtained best results respectively for AD detection and prediction, our adaptive fusion scheme enables to obtained best results by merging efficiently alteration information from both modalities.

We compared the performance of different methods applied to the whole hippocampus (see Table 6.2). The experiments showed that volume and mean of MD within a structure as the hippocampus does not provides discriminant biomarkers to detect early stages of AD. The MPBG method based on s-MRI and d-MRI obtains best results compared to the volume and the average of MD. Moreover, compared to recent methods proposed for AD detection (Arbabshirani et al., 2017) (see Table 6.4), the proposed MPBG demonstrates state-of-the-art performances for AD detection and prediction.

6.6 Conclusion

In this chapter we proposed a new multimodal approach combining a patch-based grading method and our new adaptive fusion scheme. This multimodal grading method based on s-MRI and d-MRI provides a robust hippocampal biomarker.

Indeed, the results obtained in this chapter seem confirm that the combination of structural and microstructural information enables to better track the progression of AD. However, this approach does not succeed to significantly improve classification performances compare to the use of each modality separately. This might come

from that the hippocampus is a heterogeneous structure. A better modeling of the alterations occurring into the hippocampus, at a finer grained analysis of the hippocampus structure could enable to improve the performance of hippocampal biomarkers as we will investigate it in the next chapter.

Chapter 7

Hippocampal subfields grading

Summary

7.1	Introduction	89
7.2	Materials	90
	7.2.1 MRI processing	90
	7.2.2 DTI processing	91
7.3	Experiments	92
	7.3.1 Features estimation	92
	7.3.2 Implementation	92
	7.3.3 Validation	92
	7.3.4 Statistical analyses	92
7.4	Results	93
	7.4.1 Hippocampal subfield comparisons	93
	7.4.2 Relationship with cognitive scores	94
	7.4.3 Comparison with state-of-the-art methods	95
7.5	Discussion	95
	7.5.1 Hippocampal subfield biomarkers	95
	7.5.2 Comparison with state-of-the-art methods	98
7.6	Conclusion	98

7.1 Introduction

In the previous chapters, to better capture modifications that occur during the early stage of AD, we proposed two new PBG approaches. While the first one is based on texture filters to enhance informative signal, the second embeds a multimodal analysis based on s-MRI and d-MRI. This second method aims to capture both structural and microstructural information into the same framework. Moreover, we have focused on the hippocampus alterations using a global analysis of this structure. This has been motivated by the fact that hippocampus is one of the brain structures that suffers from the most significant atrophy at the early stages of AD.

Indeed, as shown in the previous chapters, the hippocampus has been one of the most studied structures and is involved in numerous computer-aided diagnosis systems to detect AD. However, this structure is complex and not homogeneous. Hippocampus is subdivided into several subfields, each one having specific characteristics. The terminology differs across segmentation protocols (Yushkevich et al., 2015) but the most recognized definition (Lorente de Nó, 1934) mainly divides hippocampus into the subiculum, the cornu ammonis (CA1/2/3/4), and the dentate gyrus (DG). The CA1 subfield represents the biggest area in the hippocampus. It is composed by different layers called the stratum radiatum (SR), the stratum lacunosum (SL), the stratum molecular (SM), and the stratum pyramidale (SP). Furthermore, hippocampal subfields are not equally impacted by AD (Braak and Braak, 1997; Braak et al., 2006; Apostolova et al., 2006; La Joie et al., 2013; Kerchner et al., 2010, 2012; Trujillo-Estrada et al., 2014). Indeed, several MRI studies demonstrated that subfields are impacted differently according to AD stages. Postmortem, and *in vivo* imaging studies showed that CA1SR-L-M are the subfields impacted with the greatest atrophy in advanced AD (Apostolova et al., 2006; La Joie et al., 2013; Kerchner et al., 2012). Recently, it has been shown that subiculum is the earliest affected hippocampal region (Li et al., 2013; Trujillo-Estrada et al., 2014). These studies indicate that a subfield analysis of hippocampus alterations at a finer scale could provide better tool for AD detection and prediction. Finally, a recent study combining volumetric measurements and mean diffusivity of hippocampus subfields demonstrated that CA1 and subiculum are the most impacted in late AD stage (Li et al., 2013).

All these elements in addition to the results obtained in the previous chapter indicate that a multimodal analysis of hippocampal subfields using an advanced image analysis framework could provide valuable tools to improve AD detection and prediction. Consequently, in this chapter, we propose to study hippocampal subfields using s-MRI and d-MRI modalities. We study the efficiency of hippocampal subfields to improve AD detection and prediction with volume, MD and our multimodal patch-based grading method. Our results demonstrate that the study of hippocampus at finer scale improves AD prediction. Indeed, the experiments show that biomarkers based on whole hippocampus obtain best results for AD detection but biomarkers based on subiculum obtain best results for AD prediction.

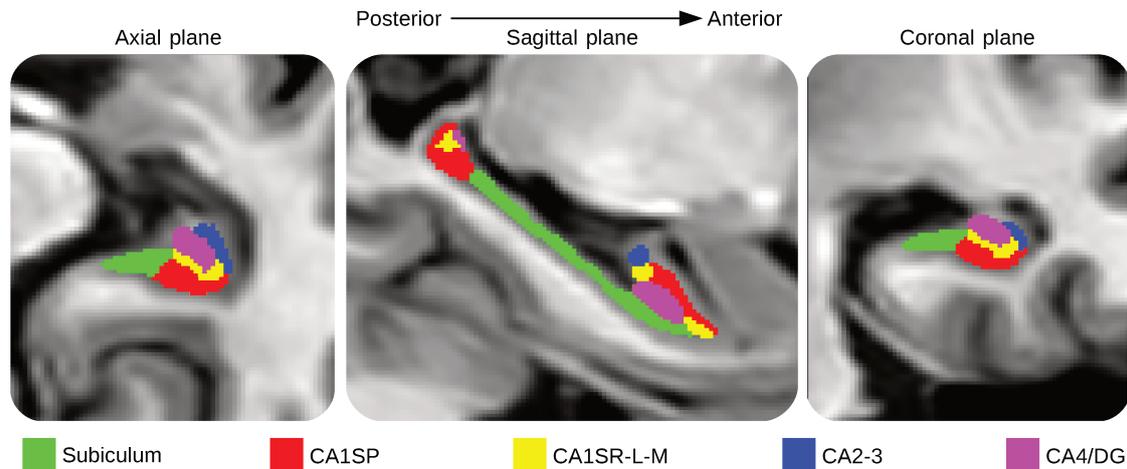


Figure 7.1: Segmentation of the hippocampal subfields. From left to right, segmentation maps of right hippocampal subfields displayed on the axial, sagittal and coronal plane.

7.2 Materials

7.2.1 MRI processing

As it is described in the Chapter 6, T1w images were processed using the volBrain system (Manjón and Coupé, 2016)¹. This system is based on an advanced pipeline providing automatic segmentation of different brain structures from T1w MRI. The preprocessing is based on: (a) a denoising step with an adaptive non-local mean filter (Manjón et al., 2010), (b) an affine registration in the MNI space (Avants et al., 2011), (c) a correction of the image inhomogeneities (Tustison et al., 2010) and (d) an intensity normalization.

Afterwards, segmentation of hippocampal subfields was performed with HIPS (Romero et al., 2017) based on a combination of non-linear registration and patch-based label fusion (Romero et al., 2016). This method uses a training library based on a dataset composed of high resolution T1w images manually labeled according to the protocol proposed by Winterburn et al. (2013) (Winterburn et al., 2013). To perform the segmentation, the ADNI images are up-sampled with a local adaptive super resolution method to fit in the training image resolution (Coupé et al., 2013). The method provides automatic segmentation of hippocampal subfields gathered into 5 labels: Subiculum, CA1SP, CA1SR-L-M, CA2-3 and CA4/DG (see Figure 7.1). At the end, the segmentation maps obtained on the up-sampled T1w images were down-sampled to fit in the MNI space resolution. All the following experiments have been carried out with images into the MNI space. Finally, an estimation of the total intra-cranial volume is performed (Manjón et al., 2014).

¹<http://volbrain.upv.es>

7.2.2 DTI processing

The pipeline described in previous chapter has been used to prepare DTI scans (see chapter 6 for further details). The preprocessing of the diffusion weighted images is based on: (a) a denoising with a LPCA filter (Manjón et al., 2013), (b) a correction of the head motion using an affine registration and (c) an affine and a non-rigid registration to the T1w MRI in the MNI space (Avants et al., 2011). Afterwards, a diffusion tensor model (Basser et al., 1994) is fitted at each voxel using Dipy library (Garyfallidis et al., 2014). To analyze microstructural modifications, the MD is estimated within each hippocampal subfield and the whole hippocampus structure with the segmentation described in the previous section.

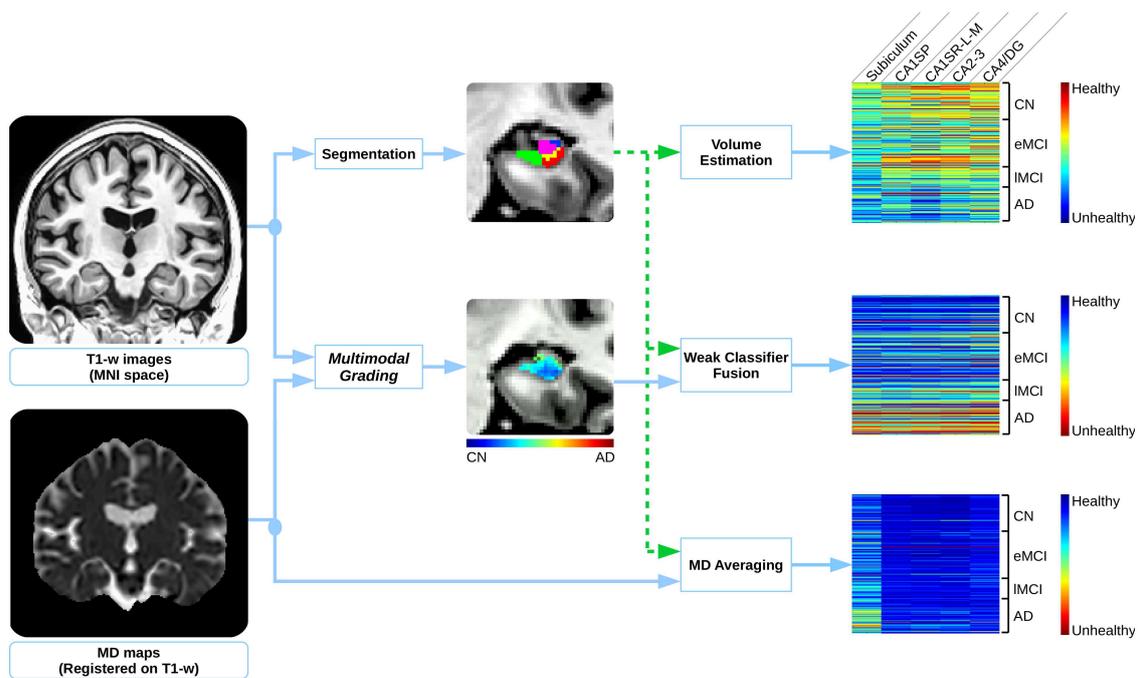


Figure 7.2: Proposed multimodal patch-based grading framework. At left, the input data: T1w images registered into the MNI space and MD maps registered on the T1w images. Data represented in this figure belongs to a CN subject. At the middle: a coronal view of hippocampal subfields segmentation on T1w, and the corresponding coronal view of a MPBG map estimated on T1w and MD. At right, the considered subfield biomarkers for all subjects under study. From top to bottom, the features are the volumes, the MPBG values, and the average of MD.

7.3 Experiments

7.3.1 Features estimation

As it has been shown in the previous chapter, the features were estimated in each hippocampal subfield and over the whole hippocampus as the union of all hippocampal subfields masks. All volumes are normalized by the total intra-cranial volume (Whitwell et al., 2001). Afterwards, we aggregate local weak classifiers of the grading map into a single feature for each considered structure (*i.e.*, hippocampal subfields and whole hippocampus). Therefore, patch-based grading features are computed by an unweighted vote of the weak classifiers using the segmentation masks (see Figure 7.2). Finally, to prevent the bias introduced as the structure alterations due to aging, all the features (*i.e.*, volume, mean of MD and MPBG) are age corrected with a linear regression based on the CN group (Dukart et al., 2011).

7.3.2 Implementation

To find the most similar patches in the training library, we use the OPAL method Giraud et al. (2016). OPAL is a fast approximate nearest neighbor patch search technique. This method enables to process each modality in about 4 seconds on a standard computer. The training library is equally composed of 37 images for both CN and AD subjects, leading to $|T| = 76$. The number of patches extracted from each the training library is $K = 160$ (*i.e.*, 80 from CN subjects and 80 from AD patients) and the patch size is $5 \times 5 \times 5$ voxels. Furthermore, we used zero normalized sum of squared differences for T1w to compute the L2 norm (see Equation (3.2)). On the other hand, d-MRI is a quantitative imaging technique. Therefore, to preserve the quantitative information, a straight sum of squared differences is used for MD in Equation (3.2),

7.3.3 Validation

To evaluate the efficiency of each considered biomarker to detect AD alterations, CN group is compared to AD patients group. In addition, to discriminate the impairment severity of MCI group, eMCI versus lMCI classification is conducted. The classification step is performed with a linear discriminant analysis (LDA) within a repeated stratified 5-fold cross-validation iterated 200 times. Mean area under the curve (AUC) and mean accuracy (ACC) are computed to compare performance for each biomarker over the 200 iterations.

7.3.4 Statistical analyses

Statistical tests have been conducted with an analysis of variances (ANOVA) procedure to determine the significance of biomarkers changes, related to the alterations

caused by AD. The results of these tests have been corrected for multiple comparisons with the Bonferroni's method. Significant changes have been tested within six comparisons (*i.e.*, CN-AD, CN-eMCI, CN-IMCI, eMCI-IMCI, eMCI-AD, and IMCI-AD). These comparisons have been achieved into each HIPP regions and with the three considered biomarkers (*i.e.*, volume, average of MD, and our newly proposed MPBG).

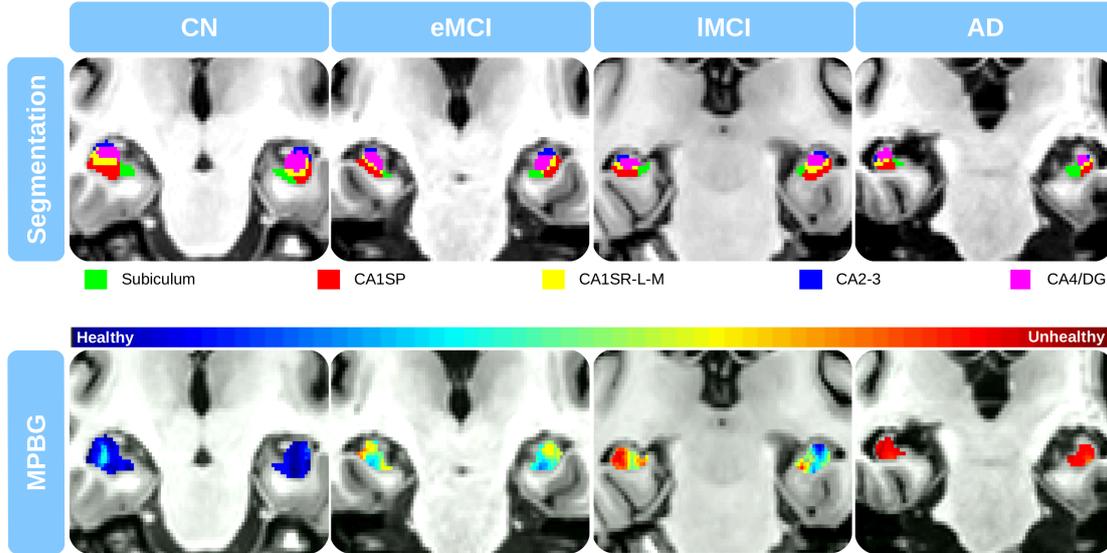


Figure 7.3: Results obtained for different severities of cognitive impairments. From top to bottom slices on the coronal plane of the segmentation maps and the fusion of T1w and MD patch-based grading with the proposed multimodal patch-based grading method. The blue and the red colors represent the healthy and altered tissues, respectively.

7.4 Results

In this section, the results are presented in two parts. First, we compare the performances in terms of AUC of each considered biomarkers within hippocampal subfields to investigate the potential of hippocampal subfields analysis to improve result of AD detection and prediction. Finally, in a last part, we compare the results of our proposed multimodal biomarker with state-of-the-art methods based on d-MRI to show the competitive performance of our approach. In addition, we provide a comparison with our texture-based method applied in different hippocampus region.

7.4.1 Hippocampal subfield comparisons

Figure 7.4 shows the distribution of volumes (A), average of MD (B) and MPBG (C) for each hippocampal subfield at each different AD stages. For each comparison a

p-value was estimated with a multi-comparison test (Hochberg and Tamhane, 1987). We can note that for all hippocampal subfields, alterations caused by the disease are related to volume and MPBG decreases while MD increase. Subiculum subfield presents the most significant differences for CN versus IMCI using volume and MD, for AD versus IMCI using MD, and for eMCI versus IMCI using MPBG. Indeed, it is the only subfield providing a p-value inferior than 0.05 for the comparison CN versus eMCI using volume, a p-value inferior than 0.01 for IMCI versus AD using MD and a p-value inferior to 0.001 to eMCI versus IMCI using MPBG, which are the most challenging comparisons. The distribution of MPBG shows a better discrimination between each group for all hippocampal subfields. Indeed, MPBG applied within CA1SP, and CA1SR-L-M provides p-values inferior than 0.01 for eMCI versus IMCI. Moreover, MPBG applied within the subiculum provides p-value inferior than 0.001 for the same comparison. Thus, MPBG enables to perform a detection of AD with each subfield with an advantage for subiculum for the comparison of eMCI versus IMCI .

To estimate the efficiency of the considered biomarkers for AD detection, we also performed a classification experiment. Figures 7.5 shows the results of two comparisons, CN versus AD (part noted A in the figure) and eMCI versus IMCI (part noted B). First, for AD diagnosis (*i.e.*, CN versus AD classification), the subfield providing the most discriminant volume is the CA1S-R-L-M with an AUC of 86.0%. Moreover, the most discriminant MD biomarker is given by the subiculum with an AUC of 88.1%. For this comparison, MD of subiculum is the only biomarker performing better results than whole hippocampus. The best results obtained by MPBG feature is provided by the CA1SP with an AUC of 92.1% followed by CA1S-R-L-M and subiculum. Second, for eMCI versus IMCI classification, the subiculum provides the best results for each considered feature. Indeed, subiculum obtained an AUC of 66.1% for the volume, 62.4% for the average of MD, and 71.8% for MPBG. Moreover, subiculum provided better results than whole hippocampus for each feature. Thus, the experiments conducted with three different biomarkers showed that the use of hippocampal subfields, especially the subiculum, enables to obtain better results for AD prediction than the whole hippocampal analysis.

7.4.2 Relationship with cognitive scores

As it has been conducted in the previous chapter, to investigate relationships between cognitive scores and MPBG values, we performed a generalized linear analysis with the following model: $MPBG = \beta_0 + \beta_1.age + \beta_2.sex + \beta_3.MMSE + \beta_4.RAVLT + \beta_5.FAQ + \beta_6.CDRSB + \beta_7.ADAS11 + \beta_8.ADAS13$. We found significant relationship of hippocampal MPBG with sex ($p < 0.01$), MMSE ($p < 0.05$) and ADAS 13 ($p < 0.01$). This results on MMSE and ADAS scores is valid for all subfields of the hippocampus. We found no specific model for a given subfield, all presented a similar pattern. These results are in line with relationships obtained between hippocampus subfields volumes and MMSE and ADAS (Khan et al., 2015).

7. Hippocampal subfields grading

Table 7.1: This table presents the p values representing the relationship of our multimodal grading hippocampal biomarkers with some meta-data (*i.e.*, sex, age, and cognitive tests). * and † indicate a p value inferior than 0.05 and 0.01, respectively.

	Sex	Age	MMSE	RAVLT	FAQ	CDRSB	ADAS11	ADAS13
Hippocampus	0.002 †	0.638	0.030 *	0.134	0.223	0.135	0.050 *	0.007 †
Subiculum	0.107	0.188	0.009 †	0.133	0.084	0.311	0.031 *	0.003 †
CA1-SP	0.001 †	0.932	0.016 *	0.081	0.269	0.149	0.109	0.016 *
CA1-SRLM	0.003 †	0.675	0.032 *	0.174	0.222	0.207	0.035 *	0.006 †
CA2/3	0.000 †	0.669	0.173	0.143	0.417	0.098	0.070	0.009 *
CA4/DG	0.003 †	0.485	0.065	0.241	0.250	0.102	0.047 *	0.010 †

7.4.3 Comparison with state-of-the-art methods

To evaluate the performance of the proposed MPBG applied on HC subfields, we compared it with state-of-the-art multimodal methods using d-MRI. To this end, we used the ACC values published by the authors. Table 7.2 shows the comparison of our proposed biomarkers within the hippocampal area providing the best results (*i.e.* the whole hippocampus and the subiculum) with the state-of-the-arts methods using similar dataset based on ADNI2.

We compare our novel hippocampal subfield biomarker with the the method described in the previous chapter. These comparisons demonstrate the relevance of MPBG biomarkers for AD detection and prediction. Indeed, our method provides similar results than the best methods with similar dataset for CN versus AD classification and provides the best results for eMCI versus IMCI classification. Moreover, the proposed MPBG method based on subiculum improves the performance for eMCI versus IMCI classification with an accuracy of 70.8%, that increases by 2% the accuracy based the whole hippocampus and over 6% compared to a connectivity network based method.

In addition, we conducted a comparison of MPBG applied within the most discriminant area for each comparison (*i.e.*, whole hippocampus and subiculum) with the texture-based grading proposed in the chapter 5 (see Table 7.3). These comparisons show that MPBG biomarkers provided better results than the texture-based grading for eMCI versus IMCI comparisons. However, texture-based grading obtains the best results for CN versus AD comparison.

7.5 Discussion

7.5.1 Hippocampal subfield biomarkers

The main contribution of this study is the multimodal analysis of hippocampal subfields. Indeed, most of the proposed biomarkers based on hippocampus focused on the whole structure or capture structural atrophy of hippocampal subfields with

Table 7.2: Comparison of our proposed MPBG biomarkers with state-of-the-arts methods based on s-MRI and d-MRI using a similar ADNI2 dataset. All results are expressed in percentage of accuracy.

Method	Subjects				Feature	Classifier	Classification ACC	
	CN	eMCI	IMCI	AD			CN/AD	eMCI/IMCI
(Nir et al., 2015)	44	74	39	23	Tractography	SVM	84.9%	n/a
(Prasad et al., 2015)	50	74	38	38	Connectivity network	SVM	78.2%	63.4%
(Zhan et al., 2015)	n/a	73	39	n/a	Connectivity network	SLG	n/a	65.0%
(Maggipinto et al., 2017)	50	22	18	50	Voxel-based	RF	87.0%	n/a
(La Rocca et al., 2018)	52	85	38	47	Connectivity network	RF	83.0%	n/a
MPBG hippocampus	62	65	34	38	Patch-based	LDA	88.1%	68.8%
MPBG Subiculum	62	65	34	38	Patch-based	LDA	86.5%	70.8%

LDA = Linear Discriminant Analysis

SLG = Sparse Logistic Regression

SVM = Support Vector Machine

RF = Random Forest

7. Hippocampal subfields grading

Table 7.3: Comparison of multimodal patch-based grading and texture-based grading (TBG) proposed in the previous chapter. The results of this comparison show that MPBG and TBG obtains similar results for eMCI versus IMCI comparison. However, for CN versus AD, TBG improve by 2 points of percentage of AUC compare to MPBG method.

Method	CN vs. AD		eMCI vs. IMCI	
	(AUC in %)	(ACC in %)	(AUC in %)	(ACC in %)
Hippocampus				
MPBG	92.1	88.1	69.5	68.8
TBG	94.1	90.2	69.2	68.1
Subiculum				
MPBG	90.9	86.5	71.8	70.8
TBG	93.5	89.0	70.6	69.7

methods based on volume despite such method are not well fitted to capture subtle alterations. The lack of work studying alterations of hippocampal subfields with advanced biomarkers could be explained by the fact that automatic segmentation of the hippocampal subfields is a complex task due to subtle borders dividing each area.

In this work, we compared the efficiency of diffusion MRI and multimodal patch-based biomarkers for AD detection and prediction over the hippocampal subfields. Comparisons based on MD, volume and multimodal patch-based biomarkers showed that the subiculum is the most discriminant structure in the earliest stage of AD providing the best results for AD prediction (see Figure 7.4 and 7.5). However, whole hippocampus structure, followed by CA1SR-L-M, obtains best results for AD detection.

These results are in accordance with literature studies based on animal model and *in vivo* imaging combining volume and MD demonstrating that subiculum is the earliest hippocampal region affected by AD (Trujillo-Estrada et al., 2014; Li et al., 2013). Moreover, postmortem studies showed that hippocampal degeneration in early stages of AD is not uniform. After the apparition of alterations in the EC, the pathology spreads to the subiculum, CA1, CA2-3 and finally the CA4 and DG subfields (Braak and Braak, 1997; Braak et al., 2006; Thal et al., 2000; Trujillo-Estrada et al., 2014). It is interesting to note that the results of our experiments using volume-based biomarkers are also coherent with the previous *in-vivo* imaging studies that analyzed the atrophy of each hippocampal subfield at advanced stage of AD. These studies showed that CA1 is the subfield impacted by the strongest atrophy (Apostolova et al., 2006; Mueller et al., 2007; La Joie et al., 2013; Carlesimo et al., 2015). Furthermore, studies using ultra-high field at 7T enabling CA1 layers discrimination showed that CA1SR-L-M is the subfields showing the greater atrophy at advanced stages of AD (Kerchner et al., 2010, 2012).

7.5.2 Comparison with state-of-the-art methods

In the past years, a large amount of studies dedicated to automatic detection of Alzheimer’s disease have been proposed (Dyrba et al., 2013; Nir et al., 2015; Li et al., 2014; Oishi et al., 2011). During our comparison, for fair comparison, we consider only methods based on the similar modalities and validated on the same ADNI2 dataset. Direct comparison with other monomodal methods applied on ADNI1 is difficult since group definition and pathological status definition are different. However, we can observe that the results obtained by the proposed method are in line with recent published results (Arbabshirani et al., 2017). Moreover, the comparison with our texture-based grading shows the interest of using biomarker based on d-MRI for the early detection of AD (see chapter 5). This comparison confirms that the higher native resolution of s-MRI may enable to capture more valuable brain modification at a late stage of the disease.

7.6 Conclusion

In this chapter, we proposed the analysis of hippocampal subfield alterations with a multimodal framework combining structural and diffusion MRI. In addition, to study tenuous modifications occurring into each hippocampal subfield, we applied our novel adaptive fusion scheme to create a new multimodal patch-based biomarker. Besides, an analysis of the hippocampal subfields with the volume, the average of MD and MBPG methods was conducted. Although CA1 is the subfields having the greater atrophy in the late stage of AD, the experiments demonstrated that whole hippocampus provides the best biomarker for AD detection while subiculum provides the best biomarker for AD prediction. Finally, we compared our novel MPBG method with the texture-based grading method proposed in the chapter 5 with the same ADNI2 dataset. The results of this comparison show that MPBG method provides better biomarker for AD prediction, even though our new texture-based method outperform the performance of patch-based grading method for AD detection.

7. Hippocampal subfields grading

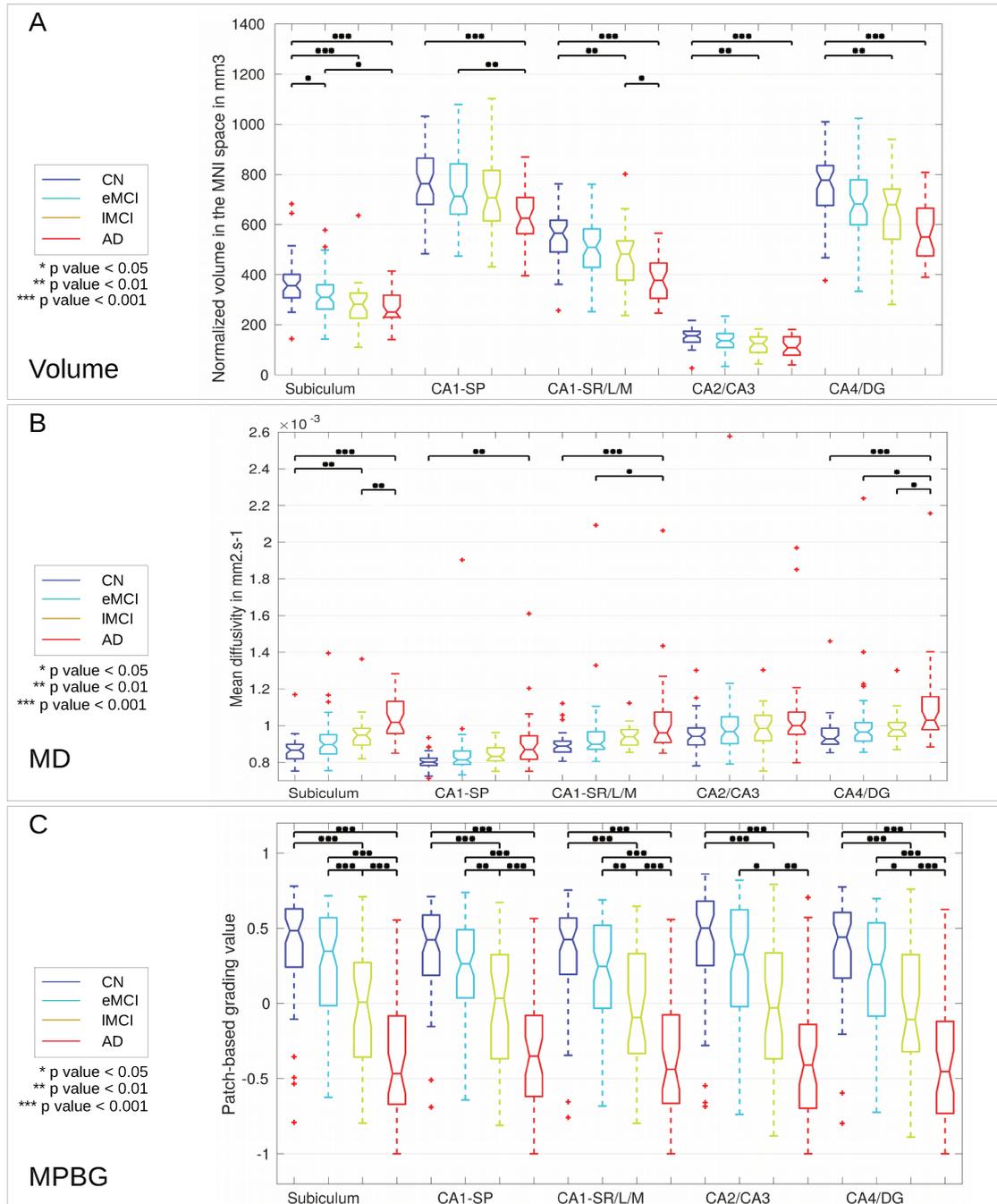


Figure 7.4: Distribution of the volume (A), MD (B) and MPBG (C) for the different considered groups. The normalized volumes are provided in mm^3 in the MNI space for each subfield, MD is the mean of MD values into each subfield in $mm^2 \cdot s^{-1}$, and MPBG is the mean patch-based grading values into each subfield. Blue, cyan, orange and red colors represent CN, eMCI, IMCI and AD subjects, respectively. Statistical tests have been performed with ANOVA procedure and corrected for multiple comparisons with the Bonferroni's method. The p-values inferior to 0.05, inferior to 0.01, and inferior to 0.001 are represented with *, **, and ***, respectively.

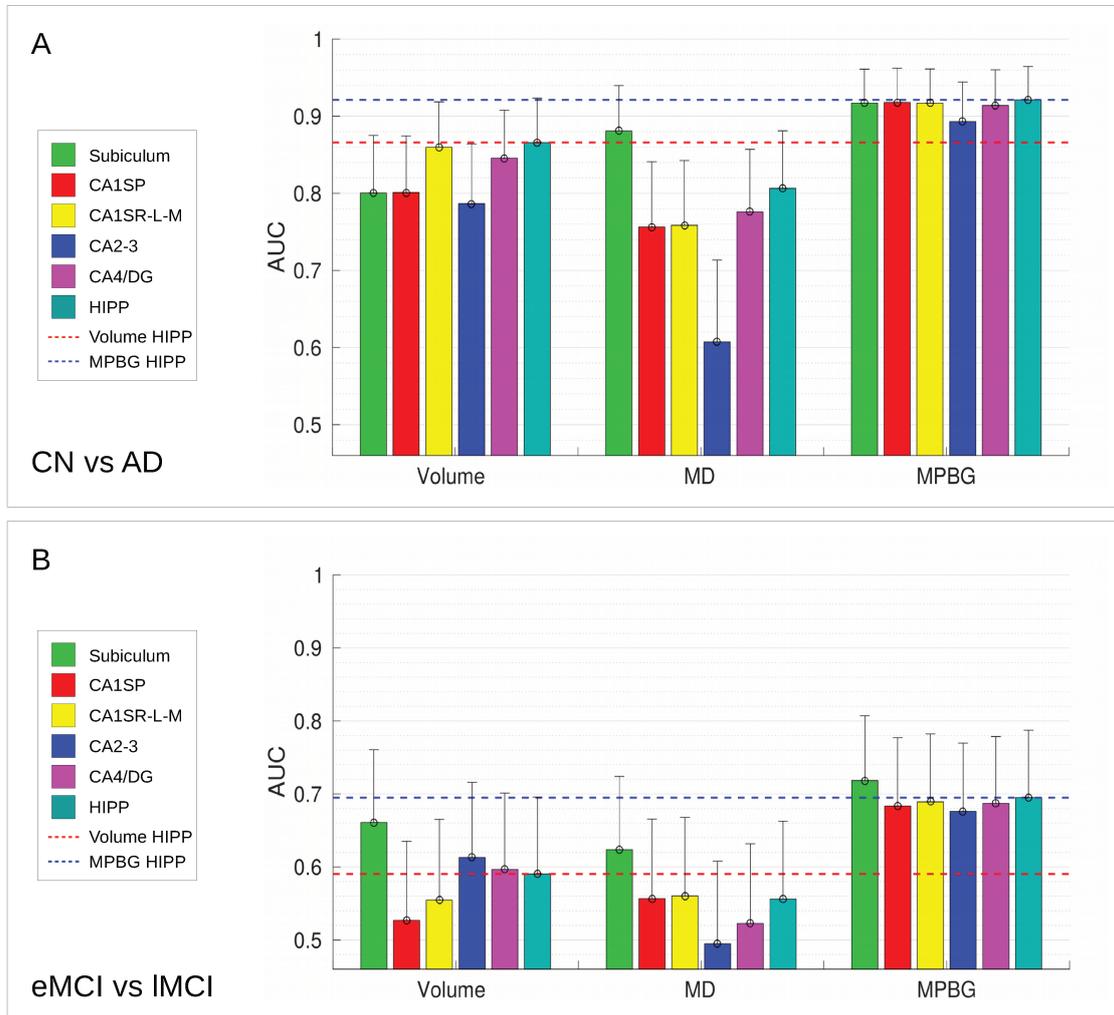


Figure 7.5: AUC computed for CN versus AD (A), eMCI versus IMCI (B) comparisons with the different considered biomarkers in each hippocampal area. Results of subfields are grouped by features (*i.e.*, the volume, the average of MD and the MPBG). Upper bounds of confidence interval are represented with vertical bars. Whole hippocampus volume biomarker provides the best results with a mean AUC of 86.6% for CN versus AD comparison, followed by the CA1S-R-L-M volume that obtains a mean AUC of 86%. Subiculum volume provides the best results for eMCI versus IMCI with a mean AUC of 66.1%. The average of MD for subiculum obtains the best results for CN versus AD and eMCI versus IMCI with a mean AUC of 88.0% and 62.4%, respectively. Whole hippocampus MPBG obtains the best results for CN versus AD with a mean AUC of 92.1%. Subiculum MPBG obtains the best results for eMCI versus IMCI comparison with a mean AUC of 71.8%. This comparison shows that subiculum is the only biomarker providing better results than whole hippocampus.

Part conclusion

In the second part of this thesis, we developed advanced hippocampus biomarkers. We first developed novel methods that we validated within the hippocampus. We proposed a texture-based grading method that obtains state-of-the-art performances for AD detection. This first contribution demonstrates that MRI intensity is not the best input and that a directional derivative filter, such as Gabor’s filter, enables to highlight informative signals. Next, we have developed a MPBG method combining s-MRI and d-MRI using our adaptive fusion scheme. The experiments conducted in this second work have shown that although s-MRI is still a good input for AD detection, the microstructural changes detected with d-MRI modality and captured with PBG methods enable a better prediction of AD progression. Moreover, our MPBG method demonstrated its ability to compute robust hippocampal biomarker that obtained best results for AD detection and prediction.

Second, we studied the hippocampus at a finer grained level. Indeed, instead of considering the hippocampus with a global approach, we studied the efficiency of hippocampal subfields. We investigated the hippocampal subfields efficiency with our novel MPBG method. Our experiments show that subiculum suffers from the most significant changes in the early stages of AD. Indeed, MPBG methods applied to the subiculum enable to increase the prediction performances compared to its application within the whole hippocampus structure. These results confirm that the most discriminant hippocampal subfields enable to obtain more effective biomarkers.

Finally, although the contributions presented in this part showed state-of-the-art results for AD detection, the prediction of subjects’ conversion to AD are still limited. Indeed, the improvement improvements achieved using new features, multiple modalities or finer-scale ROIs did not succeed to increase significantly the classification performances for AD prediction. Our main assumption is that even though our method enables to capture subtle structural modifications, a better modeling of the relationship alterations between the considered structures could provide useful information. Therefore, in the next part of this thesis, we will develop a new graph-based method that embeds inter-subjects’ similarities information and intra-subjects’ variability.

Part III:

Modeling of Alzheimer's disease signature

So far, our works have been focused on capturing anatomical alterations into key structures. However, even though the experiments conducted demonstrate an improvement for Alzheimer's disease detection, the prediction of subjects' conversion to AD remains a difficult task since the only use of inter-subjects' similarities features does not enable to improve prediction performance. Several studies have recently proposed features based on the intra-subjects' variability, suggesting that a modeling of the inter-related alterations can help to predict the dementia. Consequently, in this part, we will propose a new method based on a combination of inter-subjects' similarity and intra-subjects' variability features to better model Alzheimer's disease signature.

Introduction

In the previous parts of this thesis, we have studied advanced patch-based grading method to better capture hippocampal alterations. However, although our work yields to an improvement of performances for Alzheimer’s disease detection, these improvements show a low increase of performances for the prediction of subjects’ conversion. This might come from the global grading representation that has been used. Indeed, although the estimation of grading value distributions slightly improve the performances of prediction, it does not enable to capture complex pattern of alterations into the structures under study. We assume that a better modeling of AD signature should enable to increase performance for AD prediction.

Nowadays, the imaging-based methods developed in the literature for the prediction of the subjects’ conversion can be grouped into two categories. On the one hand, methods based on inter-subjects’ similarities have been proposed such as the PBG framework that we are studying in this thesis. On the other hand, other methods capturing an intra-subject variability by modeling the correlation of alterations of the different brain structures have also been proposed. This second group of method suggests that AD does not impact brain structures independently and thus that structures have inter-related alterations. However, the methods belonging to both groups have shown comparable classifications performances.

Consequently, we proposed to combine both inter-subjects’ similarity and intra-subjects’ variability features into a graph of structure grading method. In this part, we first apply a novel graph of structure grading within the definition of hippocampal subfields used in the previous chapter to better model the alteration of the hippocampus. This results in an increase of prediction performance compare to other hippocampal biomarkers. Second, we apply the same framework within a whole brain structure representation, leading to a considerable increase the prediction performances. Finally, we combine both anatomical scales (*i.e.*, hippocampal subfields and entire brain) into a multiple graph approach. The experiments conducted confirm our assumptions and demonstrate state-of-the-arts performances for the detection and prediction of Alzheimer’s disease.

Key-words

Intra-Subject Variability, Inter-Subject Similarity, Graph-Based Model, Brain Analysis, Hippocampal Subfields, Patch-Based Grading, Mild Cognitive Impairment, Alzheimer’s Disease Prediction

Contents

- The chapter 8 presents the general framework of the proposed graph of structure grading method. In this work, the grading values, that can be computed

from any patch-based grading method (*i.e.*, original PBG, texture-based grading, ...) are used to describe the inter-subjects' similarities and are embedded into a graph modeling to capture the relationship of alterations between different structures or sub-structures.

- The chapter 9 presents the application of our graph-based method to the hippocampal subfields definition used in the previous chapter. In this chapter, we aim at demonstrating that better modeling of hippocampus alterations enables to increase AD prediction performances. We compare patch-based grading values averaged within the whole hippocampus mask, in each different hippocampal subfields, and our graph of structure grading method. The obtained results demonstrate an increase of performance compared to other methods applied within the entire hippocampus structures.
- In the chapter 10, we apply our novel graph of structures grading to the entire brain. We compare the performance of inter-subjects' similarities, intra-subjects' variabilities features and the combination of both. The experiments conducted in this work show that intra-subjects' variability provides valuable information and its combination with inter-subjects' similarity features obtains competitive performance for AD prediction.
- The chapter 11 introduces an extension of our graph of structures grading, to combine multiple anatomical brain representation into an unified graph-based method (*i.e.*, brain structures and hippocampal subfields for instance). In this chapter, we proposed both approaches. The first one is a straightforward extension of our graph of structure grading while the second is based on a cascade of classifiers. We also study the complementarity of cognitive tests and our imaging-based method. We demonstrate in this chapter that the modeling of AD signature carried out within multiple brain structure definitions enables to increase AD detection and prediction and that our method obtains competitive performance compared to state-of-the-art approaches.

Related publications

The works of this part is based on the following papers:

Peer-reviewed international journal

[1] Multiple anatomical graph of structure grading for the prediction of Alzheimer's disease.

K. Hett, V.-T. Ta, J. V. Manjón, and P. Coupé.
NeuroImage (NI), submitted

Peer-reviewed international conferences

[2] Graph of brain structures grading for early detection of Alzheimer's disease.

K. Hett, V.-T. Ta, J. V. Manjón, and P. Coupé.
Medical Image Computing and Computer Assisted Intervention (MICCAI), 2018

[3] Graph of hippocampal subfields grading for Alzheimer's disease prediction

K. Hett, V.-T. Ta, J. V. Manjón, and P. Coupé.
Machine Learning in Medical Imaging (MLMI@MICCAI), 2018

Chapter 8

Graph of structures grading

8.1	Introduction	107
8.2	Background	107
8.3	Method	108
8.3.1	Method overview	108
8.3.2	Graph construction	109
8.3.3	Selection of discriminant graph components	110
8.4	Conclusion	110

8.1 Introduction

In the previous part of this thesis, we have proposed advanced hippocampus biomarkers to improve AD detection and prediction. The developed methods focused on a better capture of inter-subjects' similarities using structural and microstructural modifications caused by the progression of AD. However, although the proposed methods obtain state-of-the-art results for AD detection, it seems that only using methods based on inter-subjects' similarities does not enable to increase the performance of prediction for subjects' conversion. This might come from the current modeling of hippocampus alterations (*i.e.*, average and histogram representation of grading values). Indeed, it seems that the capture of alterations in specific ROIs could not provide enough information on AD. Moreover, recent works suggest that a better modeling of the alteration relationships between brain structures leads to an increase of AD prediction performances.

Consequently, in this chapter, we develop a novel framework that better captures AD signature. This method is based on the combination of the relationship of alterations through the different structures analyzed. Our model provides information of intra-subjects' variability combined with the inter-subjects' similarities computed with a PBG method that provides subtle detection of the alterations. The main assumption of our work is that AD does not cause independent alterations of structures but these alterations are inter-related. A good model of this inter-related alterations could improve the prediction performances of subjects' conversion which is still limited with our current methods. Therefore, in this work we proposed a new approach to model specific pattern of alterations caused by AD.

8.2 Background

Over the past decades, the improvement of magnetic resonance imaging (MRI) has led to the development of new imaging biomarkers (Bron et al., 2015). Many works developed biomarkers based on inter-subject similarities to detect anatomical alterations by using group-based comparison (*e.g.*, patients vs. normal controls). Some of them are based on regions of interest (ROI) to capture brain structural alterations at a large scale of analysis. The alterations of specific structures such as the cerebral cortex and hippocampus are usually captured with volume, shape, or cortical thickness (CT) measurements (Wolz et al., 2011). Other approaches proposed to study the inter-subject similarity between individuals from the same group at a voxel scale. Such methods commonly use voxel-based morphometry (VBM). VBM-based studies showed that the medial temporal lobe (MTL) is a key area to detect the first manifestations of AD (Wolz et al., 2011). Recently, more advanced methods have been designed to improve computer-aided diagnosis (Bron et al., 2015). Among them, patch-based grading (PBG) framework (Coupé et al., 2012b) proposed to better analyze inter-subject similarities. PBG uses intermediate scale between structure

and voxel and demonstrated state-of-the-art results for AD diagnosis and prognosis (Coupé et al., 2012b; Hett et al., 2017; Tong et al., 2017a).

Besides inter-subject similarity approaches, other methods proposed to capture the correlation of brain structures alterations within subjects. Indeed, although similarity-based biomarkers provide helpful tools to detect the first signs of AD, the structural alterations leading to cognitive decline are not homogeneous within a given subject. Thus, another group of biomarkers suggests that the structural changes caused by the disease may not occur at isolated areas but in several inter-related regions. Therefore, intra-subject variability feature provides relevant information. Some methods proposed to capture the relationship of spread cortical atrophy with a network-based framework (Wee et al., 2013). Other approaches estimate inter-regional correlation of brain tissues volumes (Zhou et al., 2011). A study has proposed a generic framework that embed spatial and anatomical priors within a graph modeling. This method extract inter-subject variability from different features (for instance, voxel-based and cortical thickness) and various MRI modalities (Cuingnet et al., 2013). Recently, convolutional neural network (CNN) have been used to capture relationship between anatomical structures volumes (Suk et al., 2017). Finally, some works showed that patch-based strategy can be used to model intra-subject brain alteration (Liu et al., 2014; Tong et al., 2014; Amoroso et al., 2018).

All these elements demonstrate that inter-subjects' similarity and intra-subjects' variability features provide important information on the presence of the disease. Consequently, in this work, we proposed to combine PBG framework to capture inter-subjects' similarity information with a graph-based approach to model the specific pattern of the alterations caused by AD. Indeed, we assume that a combination of inter-subjects' similarities and intra-subjects' variability method in a single modeling should provide an efficient method for AD prediction.

8.3 Method

8.3.1 Method overview

As illustrated in the Figure 8.1, our method is composed on several steps. First, a segmentation of the structures of interest is conducted on the input images. Then, a PBG approach is carried out over every segmented structures, for instance, hippocampal subfields, entire brain structures, etc. Two different factors can be detected with PBG methods: the changes caused by the normal aging (Koikkalainen et al., 2012) and the alterations caused by the progression of AD. Therefore, at each voxel, the grading values are age-corrected to avoid the bias due to normal aging. After the patch-based grading maps are age-corrected, we construct an undirected graph to model the topology of alterations caused by Alzheimer's disease. To reduce the dimensionality of the feature vector computed by our graph-based method, we

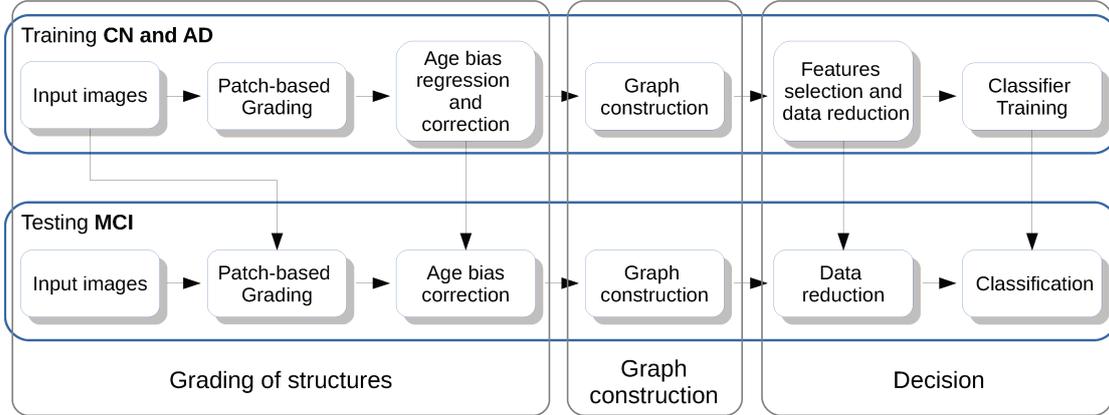


Figure 8.1: Scheme of the proposed graph of structures grading pipeline. PBG is computed using CN and AD training groups. CN group is also used to correct the bias related to age. Then, this estimation is applied to AD and MCI subjects. Afterwards, the graph is constructed, the feature selection is trained on CN and AD and is applied to CN, AD and MCI. Finally the classifier is trained with CN and AD.

use an elastic net that provides a sparse representation of the most discriminative edges and vertices of our graph, and thus enables to reduce the feature dimensionality by capturing the key regions and the key relationships between the different considered ROIs. Finally, we use only the most discriminative features of our graph as the input of a random forest method to predict the subjects' conversion.

The EN feature selection and the classifiers were trained with CN and AD. Indeed, as shown in (Tong et al., 2017a), using CN and AD to train the feature selection method and the classifier enables to better discriminate sMCI and pMCI subjects. Furthermore, it also enables to limit bias and over-fitting problem without cross-validation step.

8.3.2 Graph construction

In our graph of structure grading (GSG) method, the segmentation maps are used to fuse grading values into each ROI, and to built our graph. We define an undirected graph $G = (V, E, \Gamma, \omega)$, where $V = \{v_1, \dots, v_N\}$ is the set of vertices for the N considered brain structures and $E = V \times V$ is the set of edges. In our work, the vertices are the mean of the grading values for a given structure while the edges are based on grading distributions distances between two structures.

To this end, the probability distributions of PBG values (see Eq. 3.1) are estimated with a histogram H_v for each structure v . The number of bins is computed with the Sturge's rule (Sturges, 1926). For each vertex we assign a function $\Gamma : V \rightarrow \mathbb{R}$ defined as $\Gamma(v) = \mu_{H_v}$, where μ_{H_v} is the mean of H_v . For each edge we

assign a weight given by the function $\omega : E \rightarrow \mathbb{R}$ defined as follows:

$$\omega(v_i, v_j) = \exp(-W(H_{v_i}, H_{v_j})^2/\sigma^2) \quad (8.1)$$

where W is the Wasserstein distance with L_1 norm (Rubner et al., 2000) that showed best performance during our experiments. Wasserstein distance between two histograms is defined as the minimization of the following equation,

$$W(H_{v_i}, H_{v_j}, F) = \min_{F=\{f_{k;l}\}} \sum_{k,l} f_{k;l} d_{k;l} \quad (8.2)$$

subject to,

$$\begin{aligned} \sum_{(k) \in I} f_{k;l} &= p_k & \forall k \in I \\ \sum_{(l) \in I} f_{k;l} &= q_l & \forall l \in I \\ f_{k;l} &\geq 0 & \forall (k;l) \in J \end{aligned} \quad (8.3)$$

Where $I = \{k | 1 \leq k \leq m\}$ is the index set for bins, $H_{v_i} = \{p_k | k \in I\}$ and $H_{v_j} = \{q_k | k \in I\}$ are the two normalized histograms, $J = \{(k, l) | k \in I, l \in I\}$ is the set for flows, and $d_{k;l} = \|k - l\|_p$ is the group distance defined by a L_p norm. As described above, in our experiment we used the L_1 norm.

8.3.3 Selection of discriminant graph components

Graph representation of structure grading provides high-dimensional features. In this work we used the elastic net regression (EN) method that provides a sparse representation of the most discriminative edges and vertices, and thus enables to reduce the feature dimensionality by capturing the key structures and the key relationships between the different brain structures (see Fig. 8.1). Indeed, it has been demonstrated that combining the L_1 and L_2 norms takes into account possible inter-feature correlation while imposing sparsity (Zou and Hastie, 2005). Finally, after normalization, a concatenation of the two feature vectors is given as input of EN feature selection method, defined as the minimization of the following equation:

$$\hat{\beta} = \min_{\beta} \frac{1}{2} \|X\beta - y\|_2^2 + \rho \|\beta\|_2^2 + \lambda \|\beta\|_1 \quad (8.4)$$

Where $\hat{\beta}$ is a sparse vector that represents the regression coefficients and X is a matrix where rows correspond to the subjects and columns correspond to the features (vertices, edges or a concatenation of both).

8.4 Conclusion

In this chapter, we have developed a new graph-based model to combine inter-subjects' similarities and intra-subjects' variability. This approach models the level

of structures degradation and provides a global pattern of key structures modification. As it is described in the figure 8.1, our method is composed of three main steps: the grading of the structures of interest, that provides information about the structural alterations with an inter-subjects' similarities approach, the construction of the graph by estimating the relationship of alterations throughout the different brain structures with the distances of grading values distributions, and finally a sparse selection of the most discriminant graph components.

In the next chapters, we will study the efficiency of our novel graph-based framework with the study of hippocampal subfield alterations. Then, we will apply our new graph of structure grading with a representation of the entire brain structures. Finally, we will present an unified approach to combine our graph of structure grading within this two different anatomical scales.

Chapter 9

Graph of the hippocampal subfields

9.1	Introduction	113
9.2	Materials	113
9.3	Experiments	114
9.3.1	Preprocessing	114
9.3.2	Graph construction	114
9.3.3	Details of implementation	114
9.4	Results	115
9.4.1	Comparison of the different graph components	115
9.4.2	Comparison with state-of-the-art methods	117
9.5	Discussion	117
9.6	Conclusion	118

9.1 Introduction

In the previous part of this thesis, we capture alterations by computing global patch-based grading values into key regions. We have first studied the hippocampus structure and then compared its efficiency with the hippocampal subfields. We demonstrated that subiculum and CA1 are the best hippocampal subfields to compute global grading features. However, the results showed a low improvement in early detection performance.

We assume that a global grading value computed into specific ROIs, for instance the hippocampus or the subiculum, could limit the prediction performances since the alterations caused by AD can be inter-related. Indeed, we believe that better modeling the pattern of hippocampus alterations may improve the subjects' conversion prediction. This pattern could be based on the relationships of hippocampus alterations over the hippocampal subfields.

To confirm this assumption, we propose to apply our novel graph-based grading to hippocampal subfields. Indeed, in the previous chapter, we proposed a new graph-based framework to better model the signature of structural alterations. Our proposed method model brain alterations by combining inter-subjects' similarities and intra-subjects' variability. Therefore, the main assumption of this chapter is that our novel method can provide an efficient representation of the inter-related alterations through the hippocampus. This would results in an improvement of the performance of the prediction of subjects' conversion to AD

The experiments carried out in this chapter confirm our assumption. Indeed, the results of our experiments demonstrate that our novel approach improves patch-based grading method applied into the hippocampus by 4 percent points of accuracy and obtains state-of-the-art results compared to last advanced methods based on whole brain analysis for AD prediction.

9.2 Materials

As described in the previous chapters, data used in this work were obtained from ADNI dataset ¹. The data used in this study are all the baseline T1-w MRI of the ADNI1 phase. This dataset includes AD patients, MCI and CN subjects. The group of MCI is composed of stable MCI (sMCI) and progressive MCI (pMCI) who converted in the following 36 months after the baseline. The information of the dataset used in our work is summarized in Table 9.1.

¹<http://adni.loni.ucla.edu>

Table 9.1: Description of the dataset used in this work. Data are provided by ADNI.

Characteristic / Group	CN	sMCI	pMCI	AD
Number of subjects	213	90	126	130
Ages (years)	75.7 ± 5.0	74.9 ± 7.5	73.7 ± 7.0	74.1 ± 7.7
Sex (M/F)	108/105	58/32	68/58	64/66
MMSE	29.1 ± 1.0	27.6 ± 1.7	26.5 ± 1.6	23.5 ± 1.9

9.3 Experiments

9.3.1 Preprocessing

First, each image was preprocessed with an advanced pipeline based on: (a) a denoising step with an adaptive non-local mean filter (Manjón et al., 2010), (b) an affine registration in the MNI space (Avants et al., 2011), (c) a correction of the image inhomogeneities (Tustison et al., 2010) and (d) an intensity normalization.

Second, we used the same pipeline that has been described in the chapter 7. Consequently, the segmentation of hippocampal subfields was performed with HIPS. This method is based on a combination of non-linear registration and patch-based label fusion (Romero et al., 2017). The method provides automatic segmentation of hippocampal subfields gathered into 5 regions: Subiculum, CA1SP, CA1SR-L-M, CA2-3 and CA4/DG.

Finally, visual quality control was conducted to remove all wrong segmentations from the dataset. Moreover, to prevent any cognitive bias in the dataset, the quality control was performed without the pathological status of each subject.

9.3.2 Graph construction

Figure 9.1 represents the pipeline of the graph construction. First, the texture-based grading presented in chapter 5 is carried out over the entire hippocampus structure. Next, a histogram is computed to estimate the probability distribution of the grading values. These histograms are computed for each hippocampal subfields provided by the segmentation masks described in the previous section. Thus, here $G = (V, E, \Gamma, \omega)$, represents the graph of the hippocampal subfields where the vertices V represents alteration measures of hippocampal subfields and the edges E represent the alteration relationship between two hippocampal subfields.

9.3.3 Details of implementation

The most similar patches were extracted with a patch-match method (Giraud et al., 2016). We used the optimal parameters found in the chapter 3 for patch size and number of patch extracted. This results in hippocampus grading processed in about 1 second. Next, the age effect is corrected using linear regression estimated

on CN population. The EN method is computed with the SLEP package (Liu et al., 2009). The classifications were obtained with the random forest method (RF)¹. All features were normalized using z-score before selection and classification methods. In our experiments, we performed sMCI versus pMCI classification. However, 100 runs were performed to decrease the inner variability of RF. Mean area under curve (AUC), accuracy (ACC), balanced accuracy (BACC), sensibility (SEN), and specificity (SPE) are provided as results in Tables 9.2 and 9.3.

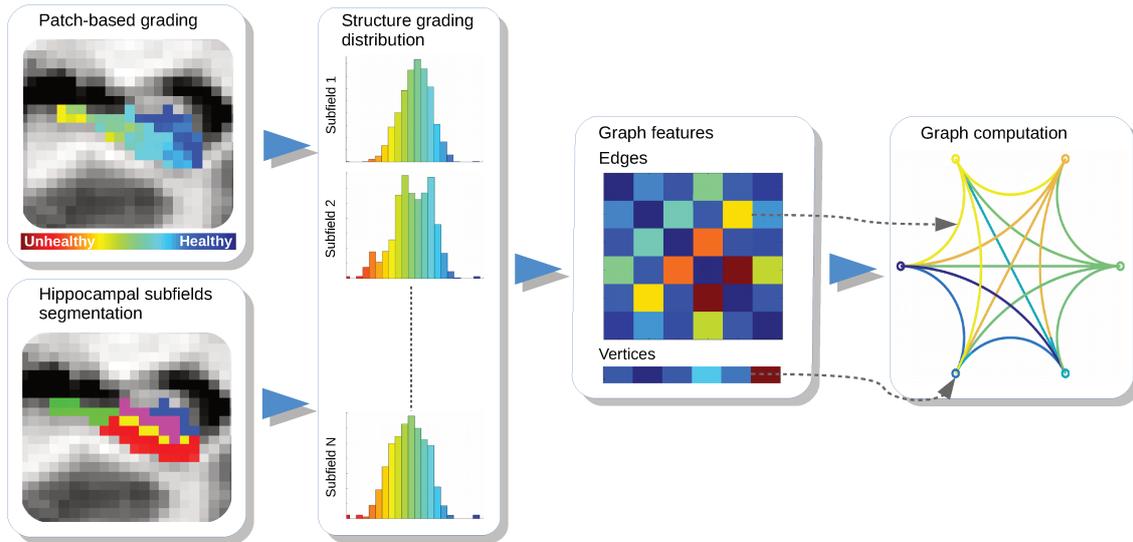


Figure 9.1: Illustration of the graph construction. From left to right, for each hippocampal region an estimation of the density probability of PBG values is computed. Next, histograms are used to built our graph of hippocampal subfields grading. Graph edges represent the distances between structure grading distribution while vertices represent the mean grading value for a given hippocampal regions (*i.e.*, hippocampus, subiculum, CA1-SP, etc.)

9.4 Results

9.4.1 Comparison of the different graph components

First, a comparison of the prediction performances was conducted with texture-based grading (TBG) previously described in the chapter 5. Thus, TBG has been applied into the whole hippocampus, the hippocampal subfields with EN selection, and our proposed graph of hippocampal subfields grading (GHSG) method (see Table 9.2). The average of TBG values into each region has been used (*i.e.*, whole hippocampus and hippocampal subfields). For hippocampal subfield features, the most relevant subfields selected are the subiculum, and the two definitions of CA1

¹<http://code.google.com/p/randomforest-matlab>

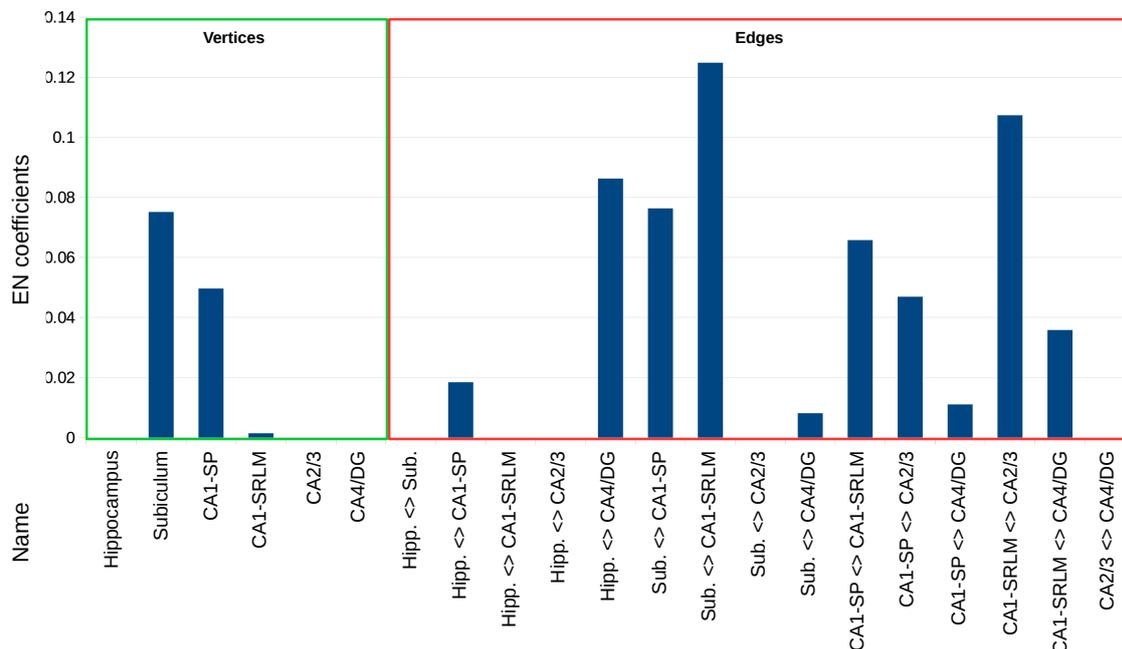


Figure 9.2: Graphic representing the coefficients estimated by EN method from the vector of features computed by our graph-based method after z-score normalization. The first six labels represent hippocampal regions and the others represent relationships of these regions. It is interesting to note that selected structures are in line with previous hippocampal subfields investigations (see chapter 7). The graph components selected are these with a EN coefficient greater than 0. In the green box, vertices of our graph. In the red box, edges of our graph.

(*i.e.*, CA1-SP and CA1-SRLM, see Figure 9.2). It is very interesting that hippocampal subfields selected by EN method are in line with previous studies which have shown CA1 and subiculum are the subfields having the most significant atrophy in late stages of AD (Kerchner et al., 2012; Trujillo-Estrada et al., 2014). TBG based on the whole hippocampus structure obtains 76.8% of AUC, 70.3% of ACC and is more specific than sensitive. Although TBG values of all hippocampal subfields (see “all” in the table 9.2) does not improve prediction performances, TBG values within subiculum, CA1-SP, and CA1-SRLM obtain 77.1% of AUC, 71.1% of ACC (see “best” in the table 9.2), and improve specificity compare to hippocampus TBG. Thus, the concatenation of mean grading values based on each hippocampal subfields selected with a EN method slightly increases the prediction performances of AD. Furthermore, our proposed graph-based method improves by 1.4 percent points of AUC and 4.4 percent points of ACC compared to the hippocampus. Our graph-based method also improves by 1.1 percent points with AUC and 3.6 percent points with ACC compared to the use of the most discriminant hippocampal subfields. Moreover, in both cases, our proposed graph-based method increases the sensibility of AD conversion.

Table 9.2: Comparison of different hippocampal PBG approaches. First, TBG applied within the whole hippocampus is provided as the baseline. Second, all hippocampal subfield features are concatenated into a single vector. Next, the best hippocampal subfield features are selected by EN method. Finally, results provided by our graph-based of hippocampal subfields grading. This comparison shows that our proposed method improves AUC, ACC, BACC, and SEN compared to other approaches. All results are given in percentage.

Methods	AUC	ACC	BACC	SEN	SPE
Hippocampus	76.8±0.2	70.3±0.0	70.6±0.0	69.0±0.0	72.2±0.0
Hipp. subfields (all)	73.9±0.2	67.1±0.0	67.9±0.0	72.2±0.0	63.5±0.0
Hipp. subfields (best)	77.1±0.2	71.1±0.4	71.4±0.4	69.5±0.6	73.2±0.5
Proposed method	78.2±0.2	74.7±0.4	74.3±0.5	77.1±0.5	71.4 ± 0.9

9.4.2 Comparison with state-of-the-art methods

Second, a comparison of our novel graph-based method based on hippocampal subfields and state-of-the-art method based on the hippocampus, using similar ADNI1 dataset, is provided in the upper part of Table 9.3. In this comparison, we included the original PBG method (Coupé et al., 2012b), a method based on multiple instance learning (Tong et al., 2014), and an advanced PBG method based on a sparse-based grading (SBG) (Tong et al., 2017a). The results demonstrate that our novel graph-based method obtains better results than all compared methods applied to the hippocampus. Indeed, to the best of our knowledge, state-of-the-art methods applied on hippocampus have obtained 71% of ACC for sMCI versus pMCI classification while our graph-based of hippocampal subfields grading obtains 74.7% of ACC.

Finally, in the lower part of Table 9.3 a comparison with state-of-the-art methods applied on the whole brain is provided. Our method is compared with a VBM approach (Moradi et al., 2015), the advanced PBG method based on a sparse-based grading (Tong et al., 2017a) and a recent deep ensemble learning method (Suk et al., 2017). This comparison shows that our novel graph of hippocampal subfields grading obtains comparable ACC and AUC than these last advanced approaches.

9.5 Discussion

In this chapter, we studied a better modeling of hippocampus alterations with the application of our graph-based framework within the hippocampal subfields. First, we studied the efficiency of a straightforward approach that consist to compute the average of grading values in each hippocampus subfield. This results lower performances compared to the average of grading values within the whole hippocampus. However, the use global grading values within each most discriminant hippocampal subfields (*i.e.*, subiculum and the two definitions of CA1) obtains similar performances than the average of grading values within the whole hippocampus. This

Table 9.3: Comparison with state-of-the-art methods based on the hippocampus region and approaches based on a whole brain analysis using similar ADNI1 dataset and the same definition of sMCI/pMCI. These results show that our proposed method obtains best results compared to methods applied within the hippocampus. Moreover, compared to approaches based on a whole brain analysis, our method obtains competitive results. All results are given in percentage.

Methods	Registration	AUC	ACC	SEN	SPE
Hippocampus					
PBG (Coupé et al., 2012b)	Affine	–	71.0	70.0	71.0
MIL (Tong et al., 2014)	Affine	–	70.4	66.5	<u>73.1</u>
SBG (Tong et al., 2017a)	Non Linear	–	69.0	–	–
Proposed method	Affine	<u>78.2</u>	<u>74.7</u>	<u>77.1</u>	71.4
Whole brain					
VBM (Moradi et al., 2015)	Non Linear	76.6	74.7	88.8	51.6
SBG (Tong et al., 2017a)	Non Linear	–	75.0	–	–
CNN (Suk et al., 2017)	Non Linear	75.4	74.8	70.9	78.8

might come from that subiculum and CA1 represent around 70% of the total volume of the hippocampus. Consequently, the average of grading values within the hippocampus work as a majority vote, and since CA1 and subiculum have the large amount of grading values. Therefore, the global hippocampus grading value is led by the grading values within subiculum and CA1. On the opposite, the use of grading values of each subfields as input of the classifier method introduce noises since the grading value of CA2/3 and CA4/DG are considered having the same relevance. Furthermore, our experiments conducted in this comparison confirm our first assumption. Indeed, our graph of hippocampal subfields grading (GHSG) method enables an improvement of the results for AD prediction.

Finally, the second comparison of the GHSG demonstrate a better AD prediction performance than state-of-the-art methods applied to hippocampus. Indeed, our method obtains better results compared to the original PBG (Coupé et al., 2012b), a patch-based grading approach based on multiple instance learning techniques (Tong et al., 2014), and an advanced grading method (Tong et al., 2017a). Moreover, compared to approaches based on whole brain analysis, our GHSG method obtains similar performances.

9.6 Conclusion

In this chapter, we have applied our graph of brain structure grading method to better capture AD signature over the hippocampal subfields. Alterations were captured with a patch-based grading framework while the relationships of alterations between the different subfields were based on histogram distances. We demonstrate

that our method improves patch-based grading methods based on hippocampus by 4 percent points for sMCI versus pMCI classification. Moreover, although the former works based on an analysis of hippocampus alterations have obtained limited performances for AD prediction compared to state-of-the-art method using a whole brain representation, our novel graph-based approach of hippocampal subfields obtains competitive results.

However, although our method applied within the hippocampus obtains competitive results, its application within an entire brain representation should increase the prediction performances. Indeed, former methods have shown that the whole brain analysis leads to an improvement of results AD prediction compared to the study of hippocampus alterations. Therefore, in the next chapter, we will propose to apply our new GSG method within an entire brain representation.

Chapter 10

Graph of the brain structures

10.1 Introduction	121
10.2 Materials	121
10.3 Experiments	122
10.3.1 Preprocessing	122
10.3.2 Graph construction	122
10.3.3 Details of implementation	123
10.4 Results	123
10.4.1 Comparisons of the different graph components	123
10.4.2 Comparisons with state-of-the-arts methods	124
10.5 Discussion	125
10.6 Conclusion	127

10.1 Introduction

In the previous chapters, we presented a new graph-based model. The proposed graph modeling encodes inter-subjects' similarities in its vertices using global PBG values and intra-subjects' variability in its edges with the distance of PBG value distributions. The application of this framework within the hippocampal subfields has demonstrated an improvement of prediction performances. Although the analysis of hippocampus alterations provides valuable biomarkers for the AD detection (Coupé et al., 2012b), former studies focused on hippocampus have obtained limited performances for the prediction of subject conversions to dementia.

However, besides these works based on the hippocampus, several studies based on the analysis of the entire brain obtained more competitive prediction results (Wee et al., 2013; Moradi et al., 2015; Tong et al., 2017a; Suk et al., 2017). Indeed, a recent study has demonstrated using the same method that the analysis carried out the whole brain instead the hippocampus enables to improve prediction of subjects' conversion significantly (Tong et al., 2017a).

Consequently, in this chapter, we proposed to apply our graph of structure grading framework to model the alterations within the entire brain structures. In our experiments, we compare the performance of intra-subject variability features (*i.e.*, the edges of our graph) with inter-subject pattern similarity features (*i.e.*, the vertices). Moreover, we demonstrate the capability of intra-subject variability features to early detect AD and show that the combination of both features improves AD prognosis. Finally, we show competitive results of our new method compared to state-of-the-art approaches.

10.2 Materials

Data use in this work were obtained from the ADNI dataset We use similar dataset than the one used in the experiments described in the chapter 5, which is based on the baseline T1w MRI of the ADNI1 phase (see Table 10.1)..

Table 10.1: Description of the ADNI dataset used in this work.

Characteristic / Group	CN	sMCI	pMCI	AD
Number of subjects	228	100	164	191
Ages (years)	75.8 ± 5.0	75.3 ± 7.2	74.2 ± 6.64	75.26 ± 7.4
Sex (M/F)	117/109	150/73	101/64	98/88
MMSE	29.05 ± 0.9	27.1 ± 2.5	26.3 ± 2.0	22.8 ± 2.9

10.3 Experiments

10.3.1 Preprocessing

The data are preprocessed using the following steps: (1) denoising using a spatially adaptive non-local means filter (Manjón et al., 2010), (2) inhomogeneity correction using N4 method (Tustison et al., 2010), (3) low-dimensional non-linear registration to MNI152 space using ANTS software (Avants et al., 2011), (4) intensity standardization, (5) segmentation using a non-local label fusion (Giraud et al., 2016) and (6) systematic error corrections (Wang et al., 2011). The patch-based multi-template segmentation was performed using 35 images manually labeled by Neuromorphometrics, Inc. ¹ follows the brain-COLOR labeling protocol composed of 134 structures.

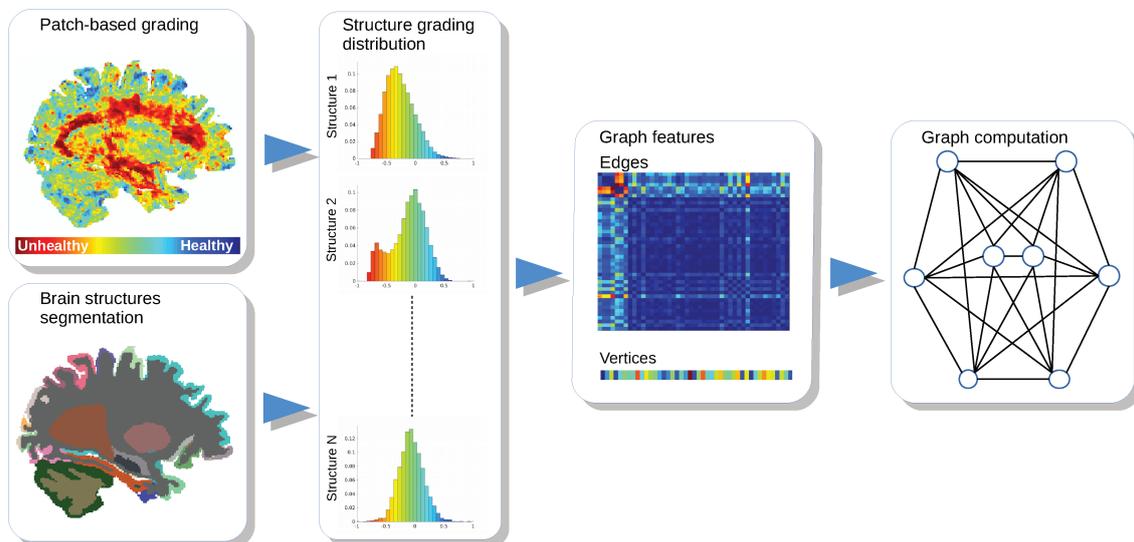


Figure 10.1: Illustration of the graph construction method. From left to right, for each segmented structure an estimation of the density probability of PBG values are computed. Then, histograms are used to build our graph of brain structure grading. Thus, 134 histograms representing each segmented brain structures are estimated. Edges are the distances between structure grading distribution while vertices are the mean grading value for a given structure (see text for more details).

10.3.2 Graph construction

Figure 10.1 represents the pipeline of the graph construction. First, the texture-based grading presented in chapter 5 is carried out over the entire brain. Next, a histogram is computed to estimate the probability distribution of grading values.

¹<http://Neuromorphometrics.com>

These histograms are computed for each brain structures provided by the segmentation masks of brain structures. Thus, here $G = (V, E, \Gamma, \omega)$, represent the graph of brain where the vertices V represent alteration measures of brain structures and the edges E represent the alteration relationship between two vertices (*i.e.*, brain structures).

10.3.3 Details of implementation

TBG were computed with the optimized patch-match method (Giraud et al., 2016) (see chapter 3). We used the optimal parameters found in the chapter 3 for patch size and number of patch extracted. This results in a whole brain grading in about 10 seconds. Age effect is corrected using linear regression estimated on CN population. The EN method is computed with the sparse learning via efficient projection package (SLEP) (Liu et al., 2009). Two classifiers were used to validate our method – support vector machine (SVM) and the random forest (RF). A linear kernel was used as SVM, which has only a soft margin parameter C . This parameter was optimized in a range of 2^i , with $i = \{-10, 9, \dots, 10\}$. RF has two parameters, the numbers of three N and the number of randomly selected features T . These two parameters was set as follows, $N = 500$ and $T = 9$. All features were normalized using z-score. In our experiments, we performed sMCI versus pMCI classification. The EN features selection and the classifiers were trained with CN and AD (see Fig. 8.1). Thus, only one run was performed for the SVM and 30 runs was performed to capture the inner variability of RF. The mean accuracy (ACC), sensibility (SEN), and specificity (SPE) over these 30 iterations are provided as results (see Table 10.2).

10.4 Results

To evaluate the improvement of performances provided by our new graph-based method, we first compare the performance of inter-subjects' similarities using the vertices of our graph, the intra-subjects' variability using the edges only, both selected with the EN regression method, and all selected GSBG features (*i.e.*, vertices and edges). Finally, we compare the performances of our new GSBG approach with the results of state-of-the-arts methods for AD prediction.

10.4.1 Comparisons of the different graph components

To investigate the results of our new GBSG method combining inter-subject pattern similarity features (*i.e.*, vertices) and intra-subject variability features (*i.e.*, the edges) several experiments were performed (see Table 10.2). The texture-based grading over the hippocampus as presented in the chapter 5 is used as baseline.

First, we estimated the performance obtained by each feature separately using SVM. Compared to hippocampus TBG, vertices showed an improvement of

Table 10.2: Classification of sMCI versus pMCI. Results obtained by inter-subject similarity features (*i.e.*, vertices), intra-subject variability features (*i.e.*, edges) and a combination of both. TBG applied over the hippocampus is provided as baseline. The results show that GBSG edges improve the accuracy and the sensibility as compared to the hippocampus PBG and GBSG vertices. Finally, GSBG provides the best results. All results are given in percentage.

Methods	Classifier	ACC	SEN	SPE
Hippocampus	SVM	71.5	72.5	70.0
Selected vertices	SVM	71.9	71.95	72.0
	RF	70.1	69.6	71.1
Selected edges	SVM	74.6	81.7	63.0
	RF	73.8	81.3	61.6
GBSG	SVM	75.8	82.3	65.0
	RF	76.5	81.7	68.0

the specificity while the accuracy and the sensibility did not change. Therefore, additional structures selected by EN did not improve results compared to use hippocampus only. On the other hand, using edges improved the accuracy and the sensibility but was less specific compared to hippocampus TBG and vertices. These results indicate that relevant information is encoded within GBSG edges.

Second, we evaluated the performance of combining vertex and edge features. GBSG provided the best results in terms of accuracy and sensibility. Moreover it improved the specificity compared to the intra-subject variability feature. Finally, we compared SVM and RF classifiers to study the stability of our framework. The results obtained with both classifier showed the same tendency. The RF provided the best results with 76.5% of accuracy.

These results obtained with two different classification methods demonstrate the complementarity of inter-subject similarity and intra-subject variability features. Indeed, both information are relevant.

10.4.2 Comparisons with state-of-the-arts methods

Afterwards, we compared our GBSG method using RF classifier with state-of-the-art methods on similar ADNI1 datasets. First, we included methods modeling inter-subject variability based on PBG within hippocampus (Coupé et al., 2012b), VBM (Moradi et al., 2015) and an advanced PBG (SBG) estimated over the entire brain (Tong et al., 2017a). Second, we included methods capturing intra-subject variability based on last deep learning framework (Suk et al., 2017), multiple instance learning (MIL) (Tong et al., 2014) and integrative network of cortical thickness abnormality (ICT) (Wee et al., 2013). We applied our GBSG on two definitions of sMCI/pMCI populations as defined in (Moradi et al., 2015) and (Tong et al., 2017a) to perform a fair comparison. Results of this comparison are presented in Table 10.3. This

comparison shows that best methods based on intra-subject or inter-subject obtained similar accuracy around 75% while our GBSG combining both reached 76.5% of accuracy. Compared to VBM on the same dataset (Moradi et al., 2015), our GBSG improved accuracy by 1.8 percent point. However, compared to SBG (Tong et al., 2017a) GBSG provided similar results on the same dataset. Finally, compared to the CNN-based method proposed in (Suk et al., 2017) our method obtained competitive performances. These results highlight the efficiency of combining intra-subject and inter-subject features.

Table 10.3: Comparison of the proposed method with state-of-the-art approaches. These results show the competitive performance of our new GBSG method that obtains the best accuracy on both definitions of sMCI/pMCI populations. All results are given in percentage.

Method	sMCI/pMCI	Area	Feature	ACC	SEN	SPE
PBG (Coupé et al., 2012b)	238/167	Hipp.	Inter	71.0	70.0	71.0
VBM (Moradi et al., 2015)	100/164	Brain	Inter	74.7	88.8	51.59
SBG (Tong et al., 2017a)	129/171	Brain	Inter	75.0	-	-
ICT (Wee et al., 2013)	111/89	Cortex	Intra	75.0	63.5	84.4
MIL (Tong et al., 2014)	238/167	MTL	Intra	72.0	69.0	74.0
CNN (Suk et al., 2017)	226/167	GM	Intra	74.8	70.9	78.8
GBSG	129/171	Brain	Inter + Intra	75.2	80.0	68.7
	100/164	Brain	Inter + Intra	76.5	81.7	68.0

10.5 Discussion

In this chapter we apply our novel graph based framework within a whole brain analysis. The comparisons of the inter-subjects' similarity and intra-subjects' variability graph components (see Table 10.2) demonstrate that capturing the information of the alterations relationship between the different brain structures enable to increase the classification performances for AD prediction. Moreover, our experiments show that most of the elements selected are graph edges, while only a few vertices representing alterations of brain structures are selected (see Figure 10.2). Indeed, only the cuneus, the medial frontal cortex, the central operculum and the postcentral gyrus are selected. Although the hippocampus is not present in the selected vertices, the hippocampus is highly present in relationship with other brain structures. This indicates that the correlation of hippocampus alteration with other brain structures is a key element to detect and predict the dementia. These results also demonstrate that the combination of both features enables a better discrimination of subject whom convert to the dementia in the following years, confirming the main assessment of the chapter 8.

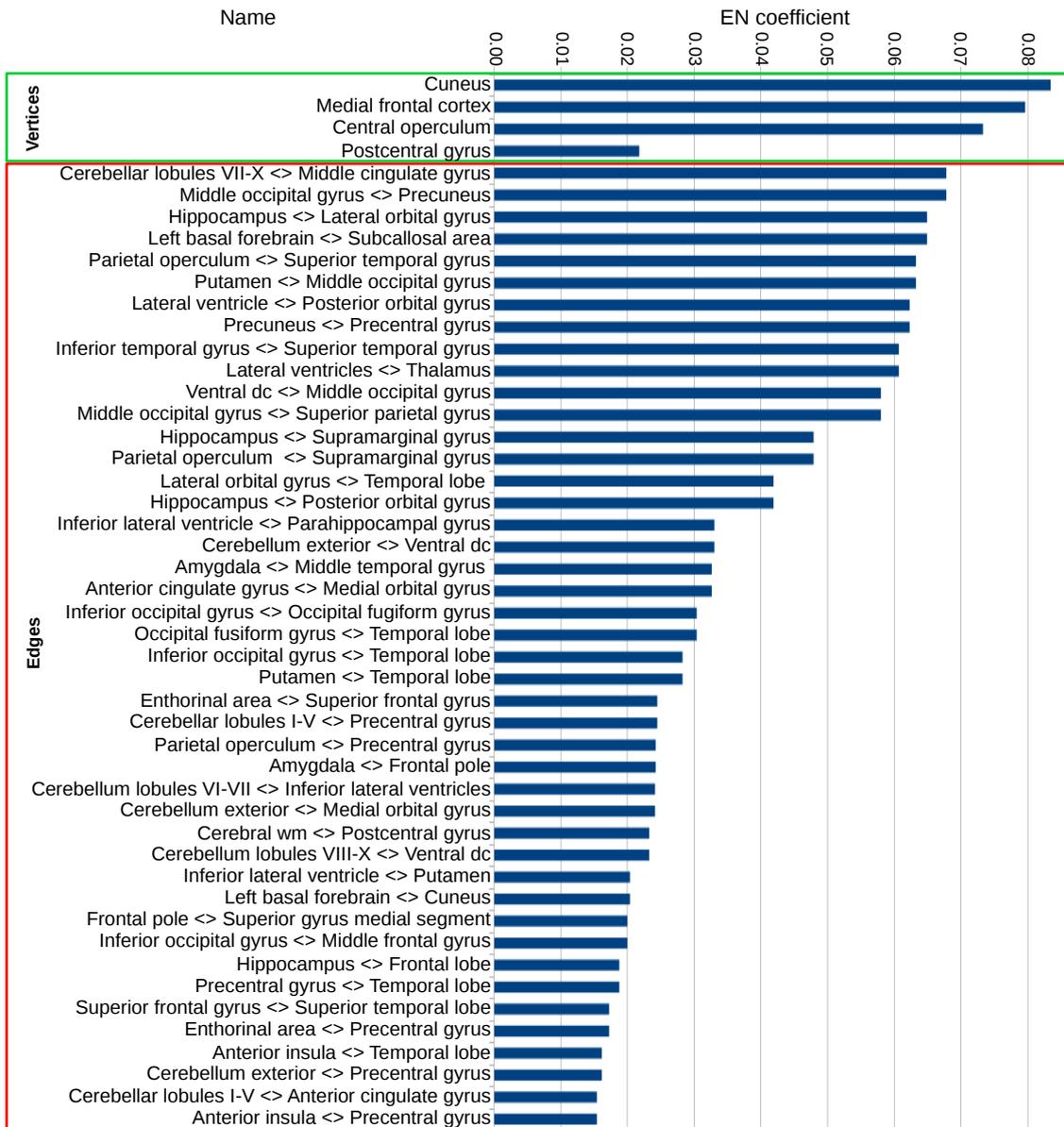


Figure 10.2: Chart of the most discriminant graph components selected by the EN regression method. The four first label correspond to the selected vertices of our graph, the others correspond to structural relationships of brain structures. This might demonstrate that intra-subjects' variability could provides more reliable informations for the prediction of AD conversion than straightforward inter-subjects' similarity informations.

Finally, the comparison of the prediction performance with state-of-the-arts methods demonstrates the efficiency of our method. Indeed, the application of our graph-based framework within the entire brain obtains competitive performances compare to last methods based on advanced VBM method (Moradi et al., 2015) and an advanced grading method (Tong et al., 2017a). Furthermore, our method obtains competitive performance compare to a deep-learning approach (Suk et al., 2017) that captures intra-subject variability of grey matter volume.

10.6 Conclusion

In this chapter, we proposed to apply our novel framework based on a promising graph of structures grading to analyze the entire brain structures. Our new method combines inter-subjects' pattern similarities and intra-subjects' variability to better detect AD alterations. The pattern similarity is estimated with a patch-based grading strategy, while the intra-subject variability between structure grading is based on graph modeling. Our experiments showed the complementarity of both information. Finally, we demonstrated that our method obtains competitive performance compared to the most advanced methods.

Chapter 11

Multiple graph of brain structures

11.1	Introduction	129
11.2	Materials	129
11.3	Experiments	130
11.3.1	Preprocessing	130
11.3.2	Fusion of graph	130
11.3.3	Combination with cognitive tests	131
11.3.4	Details of implementation	131
11.4	Results	132
11.4.1	Comparison of the different approaches	134
11.4.2	Comparison with state-of-the-art methods	134
11.4.3	Complementarity with cognitive tests	135
11.5	Discussion	136
11.5.1	Comparison of the different approaches	136
11.5.2	Comparison with state-of-the-art methods	136
11.5.3	Complementarity with cognitive tests	137
11.6	Conclusion	137

11.1 Introduction

In the third part of this thesis, we proposed a novel graph-based framework to combine inter-subjects' similarities and intra-subject' variability within an innovative graph-based model. We demonstrated the genericity of this method with its successful application to two different anatomical scale (*i.e.*, brain structure and hippocampal subfields). Indeed, our graph of structure grading applied on brain structures and the hippocampal subfields showed an improvement of performance for Alzheimer's disease prediction.

In order to combine these two anatomical scales efficiently, we proposed and compared two graph fusion methods. The first one is a straightforward extension of the graph of structures grading method described in the previous chapter. This method is similar to a work that has proposed a multi-scale analysis based on volumes of brain structures embedded into a hierarchical anatomical brain network (Zhou et al., 2011). The second approach presented in this chapter is based on a cascade of classifiers as done previously for multimodal fusion (Gray et al., 2013) or other medical imaging problematics such as the detection of other cerebral disease and multi-organ localization (Fazlollahi et al., 2014; Gauriau et al., 2015). These works have demonstrated the capability of multi-layer classifiers to improve decision precision.

In this chapter, we evaluate the performances of our two proposed approaches that combine both anatomical scales studied in the chapter 9 and 10. First, we compare two previous application of our graph of structure grading and the proposed combination of the anatomical scales. Second, we compared the best approach with state-of-the-art methods for the detection of AD and the prediction of subject conversion to AD. Finally, we study the complementarity of cognitive tests and our imaging-based method that captures brain structural alterations. All the results obtained in this chapter indicate the high potential of our proposed approach for the early detection of AD.

11.2 Materials

Data used in this work is come from the same ADNI1 dataset used in the chapter 9. The data used in this study are all the baseline T1-w MRI of the ADNI1 phase. This dataset includes AD patients, MCI and CN subjects. The information of the dataset used in our work is summarized in Table 6.1.

Table 11.1: Description of the dataset used in this work. Data are provided by ADNI.

	CN	sMCI	pMCI	AD	P value
Number	213	90	126	130	
Ages (years)	75.7 ± 5.0	74.9 ± 7.5	73.7 ± 7.0	74.1 ± 7.7	$p = 0.63^a$
Sex (M/F)	108/105	58/32	68/58	64/66	$\chi^2=5.29, p = 0.15^b$
MMSE	29.1 ± 1.0	27.6 ± 1.7	26.5 ± 1.6	23.5 ± 1.9	$p < 0.01^{a*}$
CDR-SB	3.5 ± 2.7	4.5 ± 2.3	4.8 ± 2.1	4.7 ± 1.9	$p < 0.01^{a*}$
RAVLT	45.4 ± 9.7	35.5 ± 10.2	27.7 ± 8.9	24.6 ± 7.0	$p < 0.01^{a*}$
FAQ	8.4 ± 4.4	13.3 ± 5.4	20.2 ± 6.7	30.0 ± 9.0	$p < 0.01^{a*}$
ADAS11	5.2 ± 3.0	8.1 ± 3.6	12.5 ± 4.9	20.2 ± 7.6	$p < 0.01^{a*}$
ADAS13	0.2 ± 0.9	2.3 ± 3.7	4.3 ± 4.8	14.6 ± 6.6	$p < 0.01^{a*}$

* Significant at $p < 0.05$.

^a Chi-square test (df = 3).

^b Kruskal–Wallis test (df = 3).

11.3 Experiments

11.3.1 Preprocessing

Each image has been preprocessed with the pipeline described in chapter 10 and 9. First, the data are preprocessed using the following steps: (1) denoising using a spatially adaptive non-local means filter (Manjón et al., 2010), (2) inhomogeneity correction using N4 method (Tustison et al., 2010), (3) low-dimensional non-linear registration to MNI152 space using ANTS software (Avants et al., 2011), (4) intensity standardization, (5) segmentation using a non-local label fusion (Giraud et al., 2016) and (6) systematic error corrections (Wang et al., 2011).

Brain structures have been segmented with a patch-based multi-template segmentation using 35 images manually labeled by Neuromorphometrics, Inc.¹ following the brain-COLOR labeling protocol composed of 134 structures.

Segmentation of hippocampal subfields was performed with HIPS. This method is based on a combination of non-linear registration and patch-based label fusion (Romero et al., 2017).

11.3.2 Fusion of graph

To combine this two anatomical scales (*i.e.*, brain structures and hippocampal subfields), we explore two approaches. The first one is based on a single hierarchical graph of brain structure grading while the second uses an approach based on a cascade of classifiers.

¹<http://Neuromorphometrics.com>

Hierarchical graph of structure grading (HGSG) is a straightforward extension of the graph of structures grading method described in this part. To integrate graph of hippocampal subfields grading and the graph of brain structures grading, the relationships between the whole hippocampus and each hippocampal subfield have been added in our model (see Figure 11.1). Once the graph features are computed, we use an EN regression method to select the most discriminant components of our graph. Finally, a RF is used as classifier to provide the final decision.

Multiple graphs of structure grading (MGSG) This method is based on a cascade of classifiers. In this approach, graph of brain structure grading and graph of hippocampal subfields grading are computed separately as it has been described in chapters 10 and 9 (see Figure 11.2). The EN regression method has been used to select the most discriminant features of each graph. Afterward a first layer of RF classifier is used to compute both *a posteriori* probabilities $P(Y|X_{GBSG})$ and $P(Y|X_{GHSG})$ for whole brain and hippocampal subfields, respectively. Here Y represents the pathological status of the subject under study, X_{GBSG} and X_{GHSG} represent the selected features of GBSG and GHSG models respectively. Finally, these *a posteriori* probabilities are used as input of a linear classifier to make the final decision.

11.3.3 Combination with cognitive tests

In addition of the study on the effectiveness of the proposed imaging-based biomarkers, the complementarity of our proposed method with cognitive tests has also been conducted. In this work, we consider MMSE, CDR-SB, RAVLT, FAQ, ADAS11, and ADAS13 cognitive tests. The scores are concatenated into a vector of cognitive test features and graph-based features and used as input of the final classifier.

11.3.4 Details of implementation

As shown in the previous chapters of this part, grading maps have been obtained with the texture-based grading proposed in the chapter 5. Once the grading maps are estimated, the age effect is corrected using linear regression estimated on CN population. The EN method is computed with the SLEP package (Liu et al., 2009). The classifications were obtained with a RF². In our experiments the impurity criterion used is the Gini index. RF has two parameters, the numbers of three N and the number of randomly selected features T . These two parameters was set as follows, $N = 500$ and $T = 20$ for the brain structures representation and $T = 4$ for the branch based on the hippocampal subfields. A linear discriminant analysis (LDA) classifier has been used to compute the final decision for the MGSG approach.

²<http://code.google.com/p/randomforest-matlab>

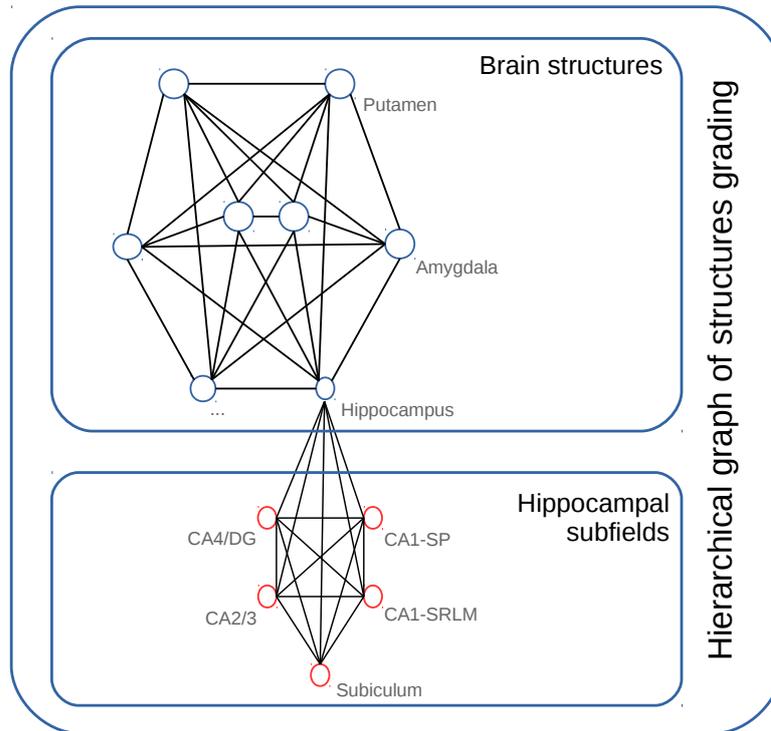


Figure 11.1: Schema of the proposed hierarchical graph of structure grading. In this approach, the two anatomical scales are embedded into a single graph-based model. The hippocampal subfield nodes are only connected with the node representing the hippocampus.

All features were normalized using z-score before selection and classification. In our experiments, we performed sMCI versus pMCI classification trained with data coming from CN and AD patients. However, 100 runs were performed to decrease the inner variability of RF. Furthermore, for CN versus AD classification, a stratified 10-fold cross-validation has been conducted 50 times. Finally, mean area under curve (AUC), accuracy (ACC), balanced accuracy (BACC), sensibility (SEN), and specificity (SPE) are provided as results in Tables 11.2, 11.3, and 11.4.

11.4 Results

In this section, to evaluate the performance of the combination of the two considered anatomical scales, the performances of HGSG and MGSG are compared. Next, we compare the best approach with state-of-the-art methods for Alzheimer’s disease classifications. Finally, we compare our imaging-based method to cognitive test scores and the combination of both.

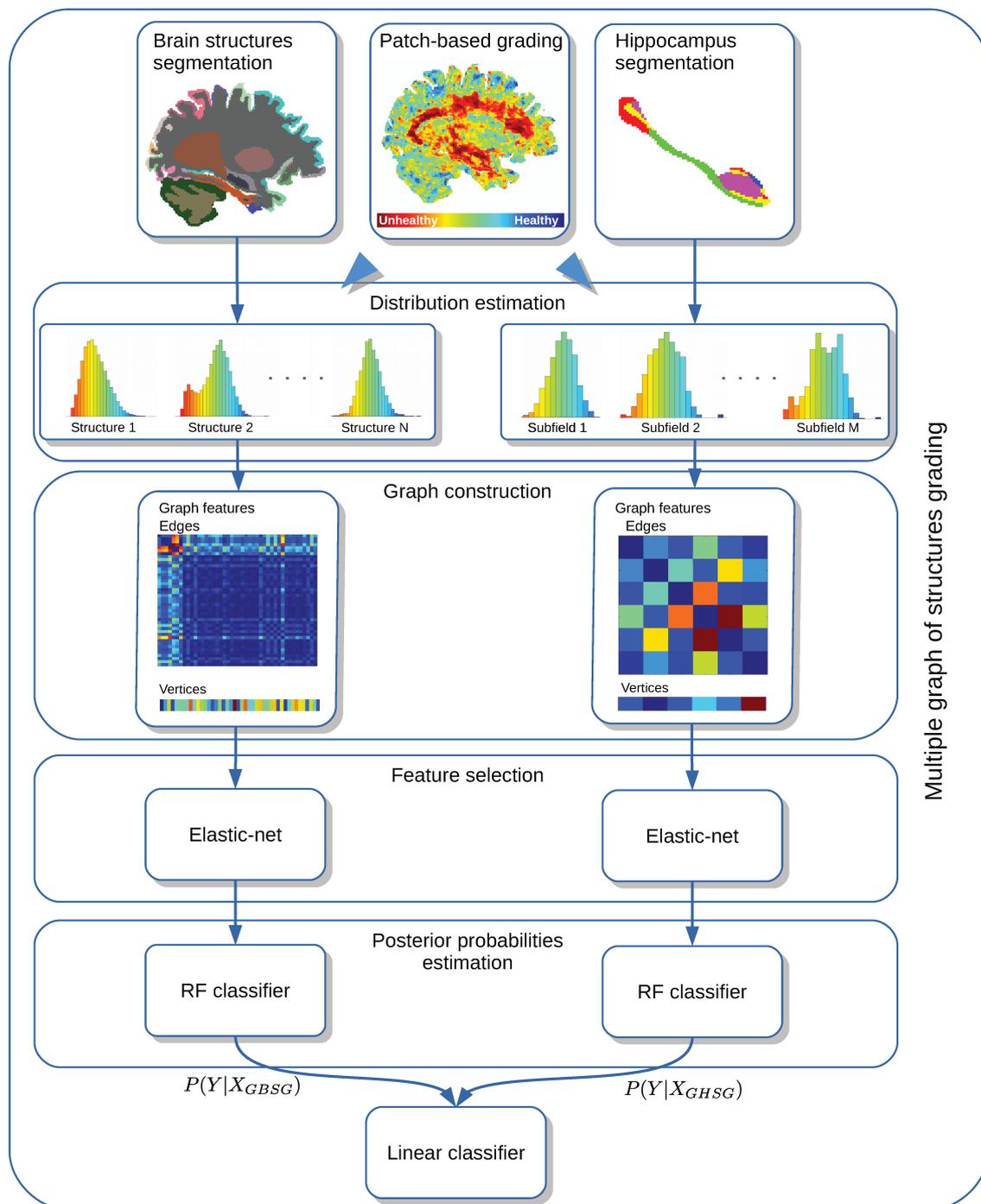


Figure 11.2: Schema of the proposed multiple graphs of structure grading method. First, each kind of segmentation map is used to aggregate grading values separately, and then a histogram is computed for each structure/subfield. Next, graphs are built and EN is used to select most discriminative graph features. A first layer of RF classifiers are computed to estimate a posterior probabilities. Finally, a linear classifier is trained with the posterior to make the final decision.

11.4.1 Comparison of the different approaches

Table 11.2: Comparisons of the different PBG approaches for sMCI versus pMCI classification task. TBG computed over the hippocampus is provided as baseline. Results show that MGSG approach improves performances in terms of AUC, ACC, BACC, SEN and SPE. All results are expressed in terms of percentage.

Methods	AUC	ACC	BACC	SEN	SPE
TBG Hipp.	76.8±0.2	70.3±0.0	70.6±0.0	69.0±0.0	72.2±0.0
GHSB	78.2±0.2	74.7±0.4	74.3 ±0.4	77.1±0.4	71.4±0.4
GBSG	79.4±0.2	75.5±0.4	75.2 ±0.4	77.6±0.4	72.6±0.4
HGSG	79.6±0.2	74.5±0.4	73.9 ±0.4	77.3±0.4	70.6±0.4
MGSG	80.6±0.2	76.0±0.4	75.7±0.4	77.8±0.4	73.6±0.4

In Table 11.2, we compare the performance of each proposed approach. TBG has been computed over the whole hippocampus and is provided as the baseline. First, we compared GHSB and GBSB within the same dataset. The results of this comparison confirm that for sMCI versus pMCI classification a whole brain analysis provide better results than considering only the hippocampus. Indeed, GBSB obtains 79.4% of AUC and 75.5% of accuracy while GHSB obtains 78.2% of AUC and 74.7% of accuracy. Second, we compare the two approaches proposed to combine both whole brain structures and hippocampal subfield. The results suggest that HGSG method does not enable to improve the performance compare to GBSB and GHSB, this method obtains 79.6% of AUC and 74.5% of ACC. However, MGSG method shows an increase of performance for each considered measures of classification. This last method obtains 80.6% of AUC and 76% of accuracy. Therefore, in the rest of the experimentations, we only consider the MGSG method.

11.4.2 Comparison with state-of-the-art methods

A comparison of MGSG results with state-of-the-art methods is provided in Table 11.3. We compare all the method using a similar ADNI1 dataset. Our graph-based method is compared with the original PBG method (Coupé et al., 2012b), a graph-based grading method (Tong et al., 2014), an ensemble grading method (Liu et al., 2012), a method based on a sparse-based grading grading (Tong et al., 2017a), a VBM method (Moradi et al., 2015) and a advanced method based on deep ensemble learning technique (Suk et al., 2017).

The results of these comparisons demonstrate the competitive performances of our MGSG method for CN versus AD and sMCI versus pMCI classifications. Indeed, our method obtains the best results with 91.6% of accuracy and 94.3% of AUC for CN versus AD which are competitive with last advanced method based on deep-learning techniques (Suk et al., 2017). Moreover, our method improves by 1 point of percentage in term of accuracy the performance of the best PBG method (Liu et

11. Multiple graph of brain structures

Table 11.3: Comparison with state-of-the-arts methods for Alzheimer’s disease classification using similar ADNI1 dataset. In addition to sMCI versus pMCI, we provided results of CN versus AD classification. All results are expressed in percentage.

Methods	Registration	Feature	CN vs. AD (ACC in %)	sMCI vs. pMCI (ACC in %)
(Coupé et al., 2012b)	Affine	Intensity	88.0	71.0
(Liu et al., 2012)	NL	GM	90.8	–
(Tong et al., 2014)	Affine	Graph	89.0	70.4
(Moradi et al., 2015)	NL	GM	–	74.7
(Tong et al., 2017a)	NL	Intensity	–	75.0
(Suk et al., 2017)	NL	GM	91.0	74.8
MGSG	Affine	Graph	91.6	76.0

GM = Grey matter

NL = Non linear

al., 2012). Furthermore, our method obtains also state-of-the-art performances for sMCI versus pMCI classification with 76.0% of accuracy and 80.6% of AUC.

11.4.3 Complementarity with cognitive tests

Table 11.4: Comparison of our graph-based approach with cognitive test scores (CS) and combination of both for AD prediction. Although, our MGSG obtains better results in terms of AUC, ACC, BACC, and SPE, the results of this comparison demonstrate the complementarity of our imaging-based method with cognitive scores. All results are expressed in percentage.

Methods	AUC	ACC	BACC	SEN	SPE
CS	78.8±0.2	74.5±0.4	72.4±0.4	84.9±0.4	60.0±0.4
MGSG	80.6±0.2	76.0±0.4	75.7±0.4	77.8±0.4	73.6±0.4
MGSG + CS	85.5±0.2	80.6±0.4	79.2±0.4	87.3±0.4	71.1±0.4

Finally, in table 11.4 we compare the results obtained using features derived from cognitive tests (CS), using our imaging-based features and the combination of both. This comparison demonstrates the superiority of our imaging-based method that improves sMCI versus pMCI classification by 1.8 points of percentage of AUC and 1.5 points of percentage of accuracy compared of using CS only. Finally, the combination of CS and MGSG features obtains 85.5% of AUC and 80.6% of accuracy which improves by around 4.9 points of percentage of AUC and 4.6% of accuracy the results of MGSG method.

11.5 Discussion

In this chapter, we proposed two approaches to achieve the combination of the whole brain structures and hippocampal subfields definition within our new graph of structure grading modeling. We compared the performances of these two approaches with the graph of structure grading applied within each structure definition. We compared our best approach to state-of-the-art methods for Alzheimer’s detection and prediction. Finally, we studied the complementarity of our proposed imaging-based method with cognitive measures.

11.5.1 Comparison of the different approaches

We first compared the results of HGSG and MSGG with the previously described GBSG and GHSG. Besides, TBG applied within the whole hippocampus structure has been provided. First, our graph of structure grading applied within hippocampal subfields definition improves prediction of Alzheimer’s disease conversion compared to other methods provided in the literature. Second, the comparison of GBSG and GHSG approaches has confirmed that the analysis of the whole brain instead specific structures such as hippocampus enables a better prediction of AD conversion (Moradi et al., 2015; Tong et al., 2017a; Suk et al., 2017).

Next, the results obtained by the straightforward extension of the graph of structure grading to combine whole brain structure and hippocampal subfields did not demonstrate an improvement of AD conversion prediction. This might come from two limitations, the augmentation of feature dimensionality and the fact that we put into the same feature space structures at different anatomical scale. To address these limitations, we proposed the MSGG method. This other approach is based on a cascade of classifiers. It helps feature selection methods to reduce high feature dimensionality efficiently. This results in an increase of AD prediction performances compared to GBSG and GHSG methods.

11.5.2 Comparison with state-of-the-art methods

We compared the results obtained with the MSGG method to state-of-the-art approaches using similar ADNI1 datasets. This comparison demonstrates the competitive performances of our novel graph-based grading method for both AD detection and prediction. The main difference with other methods come from the integration of graph-based features combining inter-subjects’ similarities and intra-subjects’ variability. Furthermore, our approach proposed for the first time to combine hippocampal subfields with brain structures for Alzheimer’s disease classification.

11.5.3 Complementarity with cognitive tests

Finally, an analysis of the complementarity of our imaging-based method with cognitive tests has been carried out. Our experiments demonstrate that our imaging-based method obtains better results for the prediction of AD. The combination of this two types of feature demonstrates a great improvement of performances compared to the use of these features separately.

11.6 Conclusion

In this chapter, we proposed a method to combine our graph-based grading framework estimated at two different anatomical scales into an unified framework. Our method proposes for the first time to combine global alterations at the brain scale and a fine-grained modifications at the hippocampal subfields scale. This framework has been initially proposed to predict the subject conversion to AD. However, the experiments conducted demonstrate state-of-the-art results for both classification tasks, the detection and the prediction of AD.

Part conclusions

In this part, we developed a new method to better model the signature of the disease. The signature of Alzheimer's disease embeds two different types of information. The first is an information based on inter-subjects' similarities that are extracted with the patch-based grading framework. The second involves an intra-subjects' variability representation of alterations captured with a graph-based modeling. Our experiments demonstrated that our modeling combining these two kinds of information enables a more effective prediction of subjects' conversion to Alzheimer's disease. the genericity of our novel methods. We applied our graph-based framework within different anatomical scales. In both cases, the experiments have shown an increase in classification performances. Thus, the state-of-the-art results for both anatomical scales demonstrate that our graph of structure grading is generic and can be applied within various representations of brain structures. Finally, we proposed a method to combine multiple anatomical scales of brain representation. We have validated our method for the combination of brain structures and hippocampal sub-fields. The experiments carried out in this part showed state-of-the-art results for the detection and prediction of AD. Moreover, our graph-based approach improves by 4 points of percentage compared with our best results obtained with the analysis of hippocampus.

Consequently, in this part, we demonstrated the need for modeling the inter-related alterations in addition of capturing subtle alterations that occurs into key ROIs and integrating features at different anatomical scales. Moreover, we have also shown the complementarity of our new imaging-based method with cognitive scores often used in clinical trials to diagnose Alzheimer's disease. This highlights the interest of using imaging-based methods in addition to cognitive scores for the early detection of Alzheimer's disease.

General conclusion and perspectives

General conclusion

In this thesis, we explored new patch-based grading methods to improve performance for detection of Alzheimer’s disease. Our works have focused on two main tasks; the detection of Alzheimer’s disease and the prediction of subjects’ conversion to this dementia. In order to increase the classification performances, we identified four key points, summarize by four questions.

The first one is “How to improve patch comparison to enhance informative signals encoded the patches?”. Indeed, patch comparison aims at capturing similar signals. Therefore, we assume that by enhancing relevant signal in the data, the patch comparison would be more accurate and would result in more reliable patch-based grading values. To answer this first question, we have proposed a new texture-based grading framework to better capture structural alterations caused by AD. Our method combines textural grading maps using a new adaptive fusion scheme. Moreover, we also have proposed an histogram-based weak classifiers aggregation approach to better discriminate early stages of AD. We have studied the optimal set of texture directions and compared our adaptive fusion to other fusion schemes. The results obtained in this work demonstrated the relevance of using textural information in combination with our novel locally adaptive fusion method. The experiments demonstrated the competitive performances of our new texture-based grading framework compared to several state-of-the-art approaches.

The second question we answered in this thesis is “How to develop an efficient multimodal approach involving d-MRI to improve performance of early detection of AD?”. So far, the patch-based grading framework has been applied to structural MRI. However, MRI techniques provide others modalities that extract different information on brain tissues. Especially, diffusion MRI seems to be a promising modality since it can capture microstructural alterations. We assumed that microstructural alterations occur before structural modifications and could lead to earlier detection of Alzheimer’s disease. To explore the use of patch-based grading method with diffusion MRI, we also proposed a new multimodal approach using our new adaptive fusion scheme. This multimodal grading method based on s-MRI and d-MRI provides a robust hippocampal biomarker. Our novel MPBG method was compared to the volume and the average of MD over the whole hippocampus. This comparison demonstrated that MPBG method improves performances for AD detection and prediction. Moreover, a comparison with state-of-the-art diffusion-based methods showed the competitive performance of our MPBG biomarker.

The third question we addressed in this thesis is “Is the study of hippocampal alterations at a finer scale, such as hippocampal subfield analysis, effective for early detection of AD?”. Indeed, hippocampus is a heterogeneous structure divided in different subfields having distinct characteristics and being unequally impacted by AD. Therefore, the study of alterations of hippocampus at a finer grained scale should improve the performance of AD detection. To this end, we proposed the analysis of hippocampal subfield alterations with our new multimodal patch-based biomarker

combining s-MRI and d-MRI. We analyzed the hippocampal subfields alterations with the volume, the average of MD and MBPG method. Our experiments showed that, although CA1 is the subfields having the greater atrophy in the late stage of AD, the whole hippocampus provides the best biomarker for AD detection while subiculum provides the best biomarker for AD prediction. Also, we compared our novel MPBG method with the proposed texture-based grading method on the same ADNI2 dataset. The results of this comparison conducted within the hippocampal subfields showed that MPBG method provides better biomarkers for AD prediction, even though our new texture-based method is better for AD detection.

Finally, the last question is “Does the combination of inter-subject similarities and intra-subject variability can increase AD prediction performance?”. To date, patch-based grading methods proposed a single feature computed by an average of the grading values within a mask. We assume that a better modeling of brain alterations could improve performances of Alzheimer’s disease detection and prediction. Consequently, we have developed a new graph-based model to embed inter-subjects’ similarities and intra-subjects’ variability. Our experiments have showed the complementarity of both information. Our method is composed of three main steps: the grading of brain structures, the construction of the graph by estimating the relationship of alterations throughout the different brain structures, and finally a sparse selection of the most discriminant graph components. First, we have applied our graph of brain structure grading method into the hippocampus. Our method is based on a graph-based representation of inter-related hippocampal subfields alterations. Second, we proposed to apply our novel framework to the brain structures. Finally, we developed a method to combine these two applications of our graph-based grading approach into a unified framework. Our method proposes for the first time to embed in a single global framework alterations at a brain scale and a hippocampal subfields scale. Our experiments demonstrated state-of-the-art results for AD detection and prediction. These results confirmed the assumption defined in the last part of this thesis. Indeed, our works demonstrated the need for modeling the inter-related alterations in addition of capturing subtle alterations that occurs into key ROIs.

The results of all contributions proposed in this thesis are summarized in the table 11.5 and compared to state-of-the-art methods presented in the first part of this manuscript. We can notice that despite the improvement proposed with TBG obtains competitive results for AD detection, the prediction have not been improved until we include intra-subjects’ variability information into our model. Furthermore, the combination of both anatomical scale enabled to obtain state-of-the-art results in both detection and prediction of AD compared with last advanced method based on deep-learning. We have not provided results of MPBG in this tabular since it has been carried out with another definition of MCI. However, MPBG obtained slightly better results than TBG for AD prediction, even though TBG obtained outperform MPBG for AD detection.

Table 11.5: Comparison of state-of-the-art methods based on similar ADNI dataset. The methods are grouped into: volume and surface-based, voxel-based morphometry, patch-based and deep-learning approaches. The results are expressed in term of detection (*i.e.* cognitively normal subject vs. Alzheimer’s disease patients) and prediction (*i.e.*, non-converted MCI vs. converted MCI) of Alzheimer’s disease. All results are expressed in percentage of accuracy for balanced group, and balanced accuracy for results based on unbalanced group of subject.

Method	Region	Feature	Classifier	Detection	Prediction
Volume and surface-based					
(Colliot et al., 2008)	Hipp	Volume	NN	84.0	-
(Querbes et al., 2009)	CC	Thickness	DT	85.0	73.0
(Gerardin et al., 2009)	Hipp	Shape	SVM	88.0	-
(Wolz et al., 2011)	Hipp	Volume	LDA	81.0	65.0
(Cuingnet et al., 2011)	Hipp	Shape	SVM	76.5*	50.0*
(Wolz et al., 2011)	CC	Thickness	LDA	81.0	56.0
(Coupé et al., 2012b)	EC	Volume	LDA	70.0	59.0
(Wee et al., 2013)	CC	Thickness	SVM	-	75.0
(Raamana et al., 2015)	CC	Thickness	MKL	89.0	64.0
Voxel-based morphometry					
(Wolz et al., 2011)	Brain	Tensor	LDA	87.0	64.0
(Abdulkadir et al., 2011)	Brain	Gray matter	SVM	87.0	-
(Cuingnet et al., 2013)	Brain	Gray matter	SVM	91.0	-
(Moradi et al., 2015)	Brain	Gray matter	LDS	-	74.7
Patch-based					
(Coupé et al., 2012b)	Hipp	Intensity	LDA	88.0	71.0
(Tong et al., 2014)	MTL	Intensity	SVM	89.0	70.0
(Tong et al., 2017a)	Hipp	Intensity	SVM	-	69.0
(Tong et al., 2017a)	Brain	Intensity	SVM	-	75.0
TBG (Hett et al., 2018d)	Hipp	Intensity	SVM	91.3	72.2
GHSg (Hett et al., 2018a)	Hipp	Intensity	RF	-	74.7
GBSG (Hett et al., 2018b)	Brain	Intensity	RF	-	75.5
MGSg (Hett et al., 2018c)	Brain	Intensity	RF	91.6	76.0
Deep learning					
(Suk et al., 2017)	Brain	Volume	CNN	91.0	74.8
(Luo et al., 2017)	Brain	Patch	CNN	83.0	-
(Liu et al., 2018)	Brain	Patch	CNN	90.5*	62.3*
(Li et al., 2018)	Brain	Patch	CNN	89.5	-

* Balanced accuracy

Perspectives

Graph of multimodal grading

The first lead of research that we would investigate is the development of graph-based model embedding different modalities such as T1, DTI, PET, etc. The first step will be the straightforward application of our multiple graph of brain structures to combine graph based on T1-weighted MRI and DTI. Indeed, the development of a multiple graph learning method is a promising path to better predict the conversion to Alzheimer's disease. Indeed, during the last years, several methods have shown that multi-modal framework enables to improve the prediction performance of Alzheimer's disease (see [Arbabshirani et al. \(2017\)](#)). However, those methods do not model the topology of the multi-modal biomarker abnormalities. Furthermore, the last graph-based methods ([Parisot et al. \(2017\)](#); [Tong et al. \(2017b\)](#)) used graph modeling to encode inter-subjects' group similarities. However, the information provided by the topology of multi-modal biomarker abnormalities could improve the classification performances. The modeling of multi-modal signature of the dementia is a challenging task, and multiple graph learning might be the best way to tackle this issue.

Deep grading framework

Recently, the deep-learning method with CNN architecture has become intensively studied, leading to many improvement in the computer recognition of natural images. These methods enable a modeling features at a higher level of abstraction. Moreover, since the similarity of two patches is computed with a linear equation, it would be possible to compute grading values with a deep-learning approach instead the current sum of squared distance combined with a non-local search. This could enable a finner modeling of patch similarity between the healthy subject and the patients suffering from Alzheimer's disease. We are currently working on this project. The first results obtained on the hippocampus area are promising.

Application of our framework to other dementia

Finally, I will explore the efficiency of our novel methods to detect and predict other brain disorders classification within a differential diagnostics framework. The differentiation between several types of neurodegenerative diseases is crucial in clinical practice. Indeed, although that Alzheimer's disease is the most common dementia for elderly people, several others brain disorders can occur during the aging, and their differential diagnosis is a challenging task. To date, methods proposed a differential diagnostics only for the detection of dementias (see [Tong et al. \(2017c\)](#); [Tolonen et al. \(2018\)](#)). However, the development of a new differential diagnostics framework that enables to better predict dementias at their first stages can help

physicians when treatment decisions have to be made. Moreover, this improvement has a significant potential for the enrichment of clinical trials.

Appendix A: Publications

Accepted manuscripts

Peer-reviewed international journals (1)

- [1] Adaptive fusion of texture-based grading for Alzheimer's disease classification.
K. Hett, V.-T. Ta, J. V. Manjón, and P. Coupé.
Computerized Medical Imaging and Graphics (COMPUT MED IMAG GRAP)

Peer-reviewed international conferences (4)

- [2] Graph of brain structures grading for early detection of Alzheimer's disease.
K. Hett, V.-T. Ta, J. V. Manjón, and P. Coupé.
Medical Image Computing and Computer Assisted Intervention (MICCAI), 2018
- [3] Graph of hippocampal subfields grading for Alzheimer's disease prediction
K. Hett, V.-T. Ta, J. V. Manjón, and P. Coupé.
Machine Learning in Medical Imaging (MLMI@MICCAI), 2018
- [4] Adaptive fusion of texture-based grading: Application to Alzheimer's disease detection.
K. Hett, V.-T. Ta, J. V. Manjón, and P. Coupé.
Patch-based Techniques in Medical Imaging (PatchMI@MICCAI), 2017
- [5] Patch-based DTI grading: Application to Alzheimer's disease classification.
K. Hett, V.-T. Ta, R. Giraud, M. Mondino, J. V. Manjón, and P. Coupé.
Patch-based Techniques in Medical Imaging (PatchMI@MICCAI), 2016

Manuscripts under consideration

Peer-reviewed international journals (2)

- [6] Multimodal hippocampal subfield grading for Alzheimer's disease classification.
K. Hett, V.-T. Ta, G. Catheline, T. Tourdias, J. V. Manjón, and P. Coupé.
Scientific Reports (Sci. Rep)
- [7] Multiple anatomical graph of structure grading for Alzheimer's disease prediction.
K. Hett, V.-T. Ta, J. V. Manjón, and P. Coupé.
NeuroImage (NI)

Appendix B: Classification methods

Linear classifier

A linear classifier achieves this by making a classification decision based on the value of a linear combination of the features. Thus, the main assessment for this group of classifier methods is that decision boundaries are linear. Such classifiers work well in the case of features are linearly separable, reaching performances comparable to non-linear classifiers with a lower computational time to train and use it.

Among linear classifier, the linear discriminant analysis (LDA) is a well-know and widely used linear classifier method is (Fisher, 1936; Lachenbruch and Goldstein, 1979). In this approach a strong assumption is made on the conditional density probabilities, x being the feature vector and y the pathological status of the subject, $p(x|y = k)$ are normally distributed with the parameters mean μ_k and the covariance matrix Σ_k , where k represents the class. LDA assumes that any class k , $\Sigma_k = \Sigma$. Thus, the linear discriminant functions is defined as follows,

$$\delta_k(x) = x^T \Sigma^{-1} \mu_k - \frac{1}{2} \mu_k^T \Sigma^{-1} \mu_k + \log \pi_k, \quad (11.1)$$

Where $\pi_k = N_k/N$ is the class-k observations. The decision is made by finding the optimal parameter with $\operatorname{argmax}_k \delta_k(x)$.

A generalization of LDA is quadratic discriminant analysis (QDA). On opposite of LDA, QDA does not assume the equality of covariant matrix Σ_k . The parameters estimations of LDA and QDA are similar, except that separate covariance matrices must be estimated for each class. Therefore, when the features dimensionality is high, the number of parameters can dramatically increase.

Support vector machine

Support vector machine (SVM) is a generalization of linear decision boundaries for classification that intend to find optimal separating hyperplanes (Cortes and Vapnik, 1995). On opposite of LDA and QDA, SVM is a non-probabilistic binary linear classifier. SVM intends to maximize the margin of the separating hyperplanes as described in Figure 11.3.

With a given vectors of pairs $(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)$, where $x_i \in R^p$ and $y_i \in \{-1, 1\}$, SVM compute the optimal separating hyperplane. To deal with the class overlap and allow for some samples to be on the wrong side of the boundary decision, slack variables $\xi = \{\xi_1, \xi_2, \dots, \xi_N\}$ are defined. These variables aim to penalize samples in the wrong boundary decision side. The equation minimization form is expressed as follows,

$$\begin{aligned} \min_{\beta, \beta_0} \frac{1}{2} \|\beta\|^2 + C \sum_{i=1}^N \xi_i \\ \text{subject to } \xi_i \geq 0, y_i(x_i^T \beta + \beta_0) \geq 1 - \xi_i \quad \forall i, \end{aligned} \quad (11.2)$$

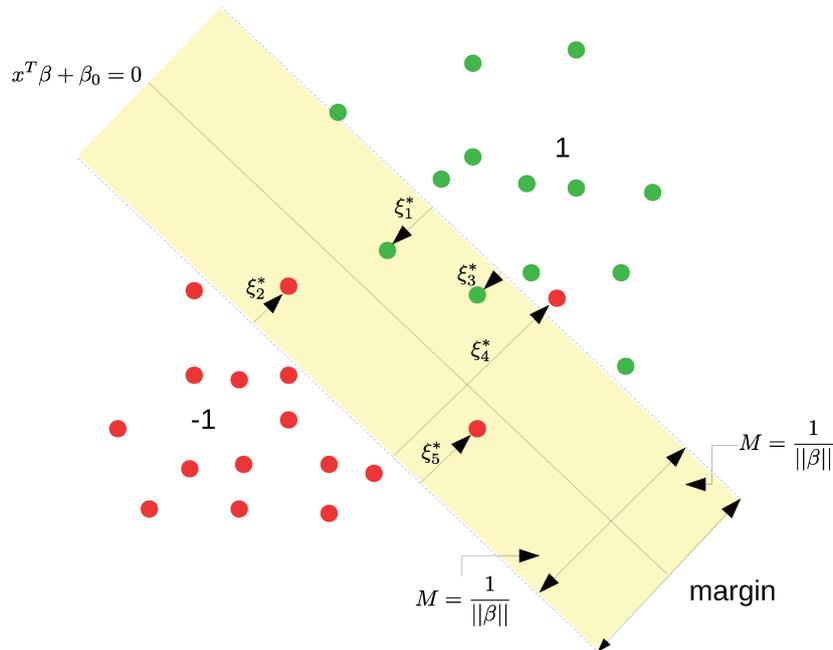


Figure 11.3: Illustration of the separating hyperplane computed with SVM in two-dimensional feature case. The decision boundary is the solid line and the dotted lines represent the maximal margin of width $2M = 2/\|\beta\|$.

Where the cost parameter C , also named soft margin parameter works as a regularization parameter that controls the trade-off between maximizing the margin and minimizing the training error.

SVM has been proposed to solve linear boundaries decision problem. However, it is possible to compute non-linear separating plane with a method named "kernel trick". The main idea is to find a relation between the features to project the feature vector that are not linearly separable in the native space of feature, into a different space where they are separable. The kernel function K should be a symmetric semi-definite function. The three popular form of kernel in the literature are:

Linear kernel: $K(x, x') = \langle x, x' \rangle$

d th-Degree polynomial: $K(x, x') = (1 + \langle x, x' \rangle)^d$

Radial basis: $K(x, x') = \exp(-\gamma \|x - x'\|^2)$

An extension of SVM has been proposed to deal with the heterogeneous vector of features. Indeed, with the complexity increase of certain model embedding a great number of different features, the use of a single kernel may in some cases project all the features in a separable space. Multiple kernel learning method (MKL)

intends to fix this issue with the use of multiple kernel extending the SVM framework (Sonnenburg et al., 2006).

Decision tree and random forest

Decision tree is a non-linear approach that aims to split the feature space with in a set of subspace and fit a simple model in each of them (Quinlan, 1986). Their concept is simple but very powerful. In Figure 11.4, we consider a two dimensional feature vector $X = (X_1, X_2)$. The first split we consider to find the threshold t_1 maximizing the separability of feature of the dimensionality X . Next, a new threshold is computed to split the subspace created by t_1 . This process continue until the stopping criterion is reached, usually it is set as the minimum node size or a certain tree depth. The algorithm intend to find the optimal split point t to separate the space R into two half subspace R_1 and R_2 , such as:

$$R_1(j, t) = \{X|X_j \leq t\} \text{ and } R_2(j, t) = \{X|X_j > t\} \quad (11.3)$$

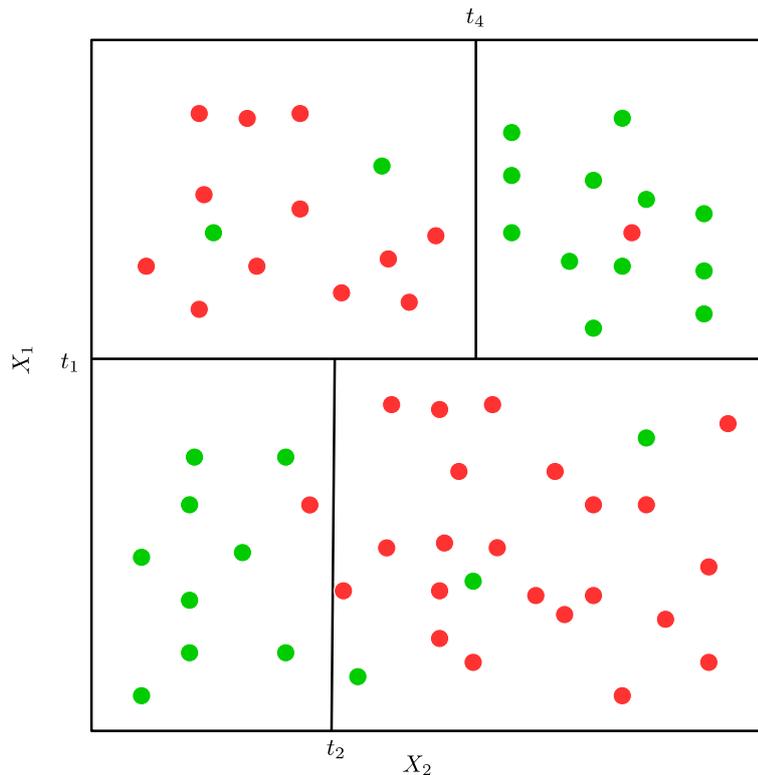


Figure 11.4: Illustration of the feature space partition conducted with the decision tree method. The values t_1 , t_2 and t_3 are the splitting points computed to partition the feature space into multiple subspaces.

The algorithm seek the splitting variable j and t that solve $\min_{j,t} G_m k$, where $G_m k$ is the impurity measure which can be defined as:

Misclassification error: $1 - \hat{p}_{mk(m)}$

Gini index: $\sum_{k=1}^K \hat{p}_{mk}(1 - \hat{p}_{mk})$

Cross-entropy: $-\sum_{k=1}^K \hat{p}_{mk} \log \hat{p}_{mk}$

Where m is the index of the terminal node, k is the class and $\hat{p}_{mk} = \frac{1}{N_m} \sum_{x_i \in R_m} I(y_i = k)$ is the proportion of of class k observations in the node m . Here N_m the number of observation, R the subspace of feature space and $I(\cdot)$ the indicator function. A pruning algorithm is often done afterward. Indeed, the complexity of the tree is defined as the number of split. The reduction of this complexity can prevent for over fitting the training data. Finally, the classification of the observation is made with $k(m) = \operatorname{argmax}_k \hat{p}_{mk}$.

To reduce variance of an estimated prediction function, bagging and bootstrap aggregation techniques have been proposed. Random forest (RF) (Breiman, 2001) is a modification of bagging that builds a large set of de-correlated decision trees, and use a majority voting to make the final decision. RF became very popular since their are simple to train and use and have high abilities to deal with heterogeneous data.

For each tree in the RF, a bootstrap sample of the training data is selected. On opposite of decision tree, each splitting step involve a random process that select l feature from the vector of feature and pick the best feature and split point among those l features.

Neural network

A neural network is a nonlinear model that consists of units (neurons), arranged in layers (see Figure 11.5). The central idea is to extract linear combinations of the inputs as derived features and model the target as a nonlinear function of these features. The networks are defined feed-forward way as follows: a unit feeds its output to all the units on the next layer. Each unit takes an input, applies a function to it and then passes the output on to the next layer (LeCun et al., 2015).

In other words, for a 2-class classification, there are two units at the end of the network, each one modeling the probability of the class that the unit represents. Hence, there are K units for a K -class classification network, the K units measure Y_k , with $k = 1, \dots, K$ each having a value between 0 and 1. For a single layer network as illustrated in Figure 11.5, the derived features Z_m , named hidden layer, are created from linear combinations of the inputs and the target Y_k is modeled as a function of linear combination of the Z_m , such as:

$$\begin{aligned} Z_m &= \sigma(\alpha_{0m} + \alpha_m^T X), m = 1, \dots, M \\ f_k(X) &= g_k(\beta_{0k} + \beta_k^T Z), k = 1, \dots, K \end{aligned} \tag{11.4}$$

where $Z = (Z_1, Z_2, \dots, Z_M)$. The function $\sigma(\cdot)$ is named activation function. The first activation function was chosen to be a sigmoid, or hyperbolic tangent. However, the recent neural network usually choose to a rectified linear unit (ReLU) that offer practical benefits compare to the sigmoid function. The main benefit of the ReLU come from its sparsity and by reducing likelihood of the gradient to vanish. Indeed, the training phase of neural network induces a backward propagation that adjusts the parameters of each layers fitting the network to the classification problem. Finally, the probability of each class k is usually computed with a softmax function g_k defined as:

$$g_k(T) = \frac{e^{T_k}}{\sum_{l=1}^K e^{T_l}} \quad (11.5)$$

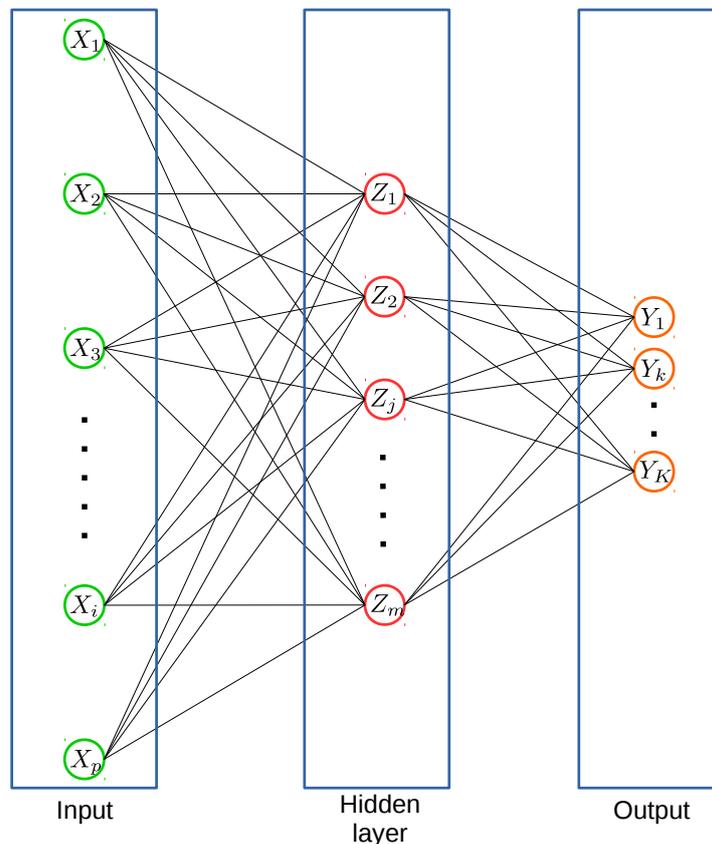


Figure 11.5: Illustration of a single hidden layer neural network. X_i represents the input i (*i.e.* the feature i), Z_j represents the neurons j in the hidden layer of neurons and Y_k is the target output of class k .

With the recent development of deep-learning methods, neural-network has become highly popular outperforming numerous state-of-the-art methods in pattern recognition fields. However, in the neuroimaging fields, such approaches have barely obtained competitive performance compare to more simple methods. The main lim-

itations of these methods come from their great number to optimize, needing large training libraries.

Bibliography

- [Abdulkadir et al. 2011] ABDULKADIR, Ahmed ; MORTAMET, Bénédicte ; VEMURI, Prashanthi ; JACK JR, Clifford R. ; KRUEGER, Gunnar ; KLÖPPEL, Stefan ; INITIATIVE, Alzheimer’s Disease N. et al. : Effects of hardware heterogeneity on the performance of SVM Alzheimer’s disease classifier. In : *Neuroimage* 58 (2011), n. 3, pp. 785–792 [30](#), [35](#), [142](#)
- [Acharya et al. 2012] ACHARYA, U R. ; SREE, S V. ; ALVIN, Ang Peng C. ; SURI, Jasjit S. : Use of principal component analysis for automatic classification of epileptic EEG activities in wavelet framework. In : *Expert Systems with Applications* 39 (2012), n. 10, pp. 9072–9078 [20](#)
- [Allen 1974] ALLEN, David M. : The relationship between variable selection and data augmentation and a method for prediction. In : *Technometrics* 16 (1974), n. 1, pp. 125–127 [23](#)
- [Alzheimer’s 2015] ALZHEIMER’S, Association : 2015 Alzheimer’s disease facts and figures. In : *Alzheimer’s & dementia: the journal of the Alzheimer’s Association* 11 (2015), n. 3, pp. 332 [3](#)
- [Amoroso et al. 2018] AMOROSO, Nicola ; LA ROCCA, Marianna ; BRUNO, Stefania ; MAGGIPINTO, Tommaso ; MONACO, Alfonso ; BELLOTTI, Roberto ; TANGARO, Sabina : Multiplex Networks for early diagnosis of Alzheimer’s disease. In : *Frontiers in Aging Neuroscience* 10 (2018), pp. 365 [108](#)
- [Anand and Sahambi 2010] ANAND, C S. ; SAHAMBI, Jyotinder S. : Wavelet domain non-linear filtering for MRI denoising. In : *Magnetic Resonance Imaging* 28 (2010), n. 6, pp. 842–861 [16](#)
- [Anoop et al. 2010] ANOOP, A ; SINGH, Pradeep K. ; JACOB, Reeba S. ; MAJI, Samir K. : CSF biomarkers for Alzheimer’s disease diagnosis. In : *International journal of Alzheimer’s disease* 2010 (2010) [6](#)
- [Apostolova et al. 2006] APOSTOLOVA, Liana G. ; DUTTON, Rebecca A. ; DINOVI, Ivo D. ; HAYASHI, Kiralee M. ; TOGA, Arthur W. ; CUMMINGS, Jeffrey L. ; THOMPSON, Paul M. : Conversion of mild cognitive impairment to Alzheimer disease predicted by hippocampal atrophy maps. In : *Archives of neurology* 63 (2006), n. 5, pp. 693–699 [89](#), [97](#)
- [Arbabshirani et al. 2017] ARBABSHIRANI, Mohammad R. ; PLIS, Sergey ; SUI, Jing ; CALHOUN, Vince D. : Single subject prediction of brain disorders in

BIBLIOGRAPHY

- neuroimaging: Promises and pitfalls. In : *NeuroImage* 145 (2017), pp. 137–165
15, 86, 98, 143
- [Ashburner 2007] ASHBURNER, John : A fast diffeomorphic image registration algorithm. In : *Neuroimage* 38 (2007), n. 1, pp. 95–113 18
- [Ashburner and Friston 2000] ASHBURNER, John ; FRISTON, Karl J. : Voxel-based morphometry—the methods. In : *Neuroimage* 11 (2000), n. 6, pp. 805–821
20, 21, 29
- [Ashburner and Friston 2005] ASHBURNER, John ; FRISTON, Karl J. : Unified segmentation. In : *Neuroimage* 26 (2005), n. 3, pp. 839–851 26
- [Ashburner et al. 2000] ASHBURNER, John ; GOOD, Catriona ; FRISTON, Karl J. : Tensor based morphometry. In : *NeuroImage* 5 (2000), n. 11, pp. S465 30
- [Association et al. 2008] ASSOCIATION, Alzheimer’s et al. : 2008 Alzheimer’s disease facts and figures. In : *Alzheimer’s & dementia: the journal of the Alzheimer’s Association* 4 (2008), n. 2, pp. 110 3
- [Avants et al. 2011] AVANTS, Brian B. ; TUSTISON, Nicholas J. ; SONG, Gang ; COOK, Philip A. ; KLEIN, Arno ; GEE, James C. : A reproducible evaluation of ANTs similarity metric performance in brain image registration. In : *Neuroimage* 54 (2011), n. 3, pp. 2033–2044 18, 65, 79, 80, 90, 91, 114, 122, 130
- [Axel et al. 1987] AXEL, Leon ; COSTANTINI, Jay ; LISTERUD, John : Intensity correction in surface-coil MR imaging. In : *American Journal of Roentgenology* 148 (1987), n. 2, pp. 418–420 17
- [Babalola et al. 2009] BABALOLA, Kolawole O. ; PATENAUDE, Brian ; ALJABAR, Paul ; SCHNABEL, Julia ; KENNEDY, David ; CRUM, William ; SMITH, Stephen ; COOTES, Tim ; JENKINSON, Mark ; RUECKERT, Daniel : An evaluation of four automatic methods of segmenting the subcortical structures in the brain. In : *Neuroimage* 47 (2009), n. 4, pp. 1435–1447 19
- [Bai et al. 2015] BAI, Wenjia ; SHI, Wenzhe ; LEDIG, Christian ; RUECKERT, Daniel : Multi-atlas segmentation with augmented features for cardiac MR images. In : *Medical image analysis* 19 (2015), n. 1, pp. 98–109 59
- [Barnes et al. 2009] BARNES, Connelly ; SHECHTMAN, Eli ; FINKELSTEIN, Adam ; GOLDMAN, Dan : PatchMatch: A randomized correspondence algorithm for structural image editing. In : *ACM Transactions on Graphics-TOG* 28 (2009), n. 3, pp. 24 41, 57
- [Baron et al. 2001] BARON, JC ; CHETELAT, G ; DESGRANGES, B ; PERCHEY, G ; LANDEAU, B ; DE LA SAYETTE, V ; EUSTACHE, F : In vivo mapping of

- gray matter loss with voxel-based morphometry in mild Alzheimer's disease. In : *Neuroimage* 14 (2001), n. 2, pp. 298–309 [30](#)
- [Basser et al. 1994] BASSER, Peter J. ; MATTIELLO, James ; LEBIHAN, Denis : MR diffusion tensor spectroscopy and imaging. In : *Biophysical journal* 66 (1994), n. 1, pp. 259–267 [80](#), [91](#)
- [Black and Sapiro 1999] BLACK, Michael J. ; SAPIRO, Guillermo : Edges as outliers: Anisotropic smoothing using local image statistics. In : *International Conference on Scale-Space Theories in Computer Vision* Springer (Veranst.), 1999, pp. 259–270 [16](#)
- [Bobinski et al. 1999] BOBINSKI, Maciek ; DE LEON, Mony J. ; CONVIT, Antonio ; DE SANTI, Susan ; WEGIEL, Jerzy ; TARSHISH, Chaim Y. ; SAINT LOUIS, LA ; WISNIEWSKI, Henryk M. : MRI of entorhinal cortex in mild Alzheimer's disease. In : *The Lancet* 353 (1999), n. 9146, pp. 38–40 [3](#), [26](#)
- [Boccardi et al. 2015] BOCCARDI, Marina ; BOCCHETTA, Martina ; MORENCY, Félix C ; COLLINS, D L. ; NISHIKAWA, Masami ; GANZOLA, Rossana ; GROTHE, Michel J. ; WOLF, Dominik ; REDOLFI, Alberto ; PIEVANI, Michela : Training labels for hippocampal segmentation based on the EADC-ADNI harmonized hippocampal protocol. In : *Alzheimer's & Dementia* 11 (2015), n. 2, pp. 175–183 [66](#)
- [Braak and Braak 1997] BRAAK, E ; BRAAK, H : Alzheimer's disease: transiently developing dendritic changes in pyramidal cells of sector CA1 of the Ammon's horn. In : *Acta neuropathologica* 93 (1997), n. 4, pp. 323–325 [89](#), [97](#)
- [Braak et al. 2006] BRAAK, Heiko ; ALAFUZOFF, Irina ; ARZBERGER, Thomas ; KRETZSCHMAR, Hans ; DEL TREDICI, Kelly : Staging of Alzheimer disease-associated neurofibrillary pathology using paraffin sections and immunocytochemistry. In : *Acta neuropathologica* 112 (2006), n. 4, pp. 389–404 [89](#), [97](#)
- [Brechtbühler et al. 1995] BRECHBÜHLER, Ch ; GERIG, Guido ; KÜBLER, Olaf : Parametrization of closed surfaces for 3-D shape description. In : *Computer vision and image understanding* 61 (1995), n. 2, pp. 154–170 [27](#)
- [Breiman 2001] BREIMAN, Leo : Random forests. In : *Machine learning* 45 (2001), n. 1, pp. 5–32 [21](#), [151](#)
- [Bron et al. 2015] BRON, Esther E. ; SMITS, Marion ; VAN DER FLIER, Wiesje M. ; VRENKEN, Hugo ; BARKHOF, Frederik ; SCHELTENS, Philip ; PAPMA, Janne M. ; STEKETEE, Rebecca M. ; ORELLANA, Carolina M. ; MEIJBOOM, Rozanna : Standardized evaluation of algorithms for computer-aided diagnosis of dementia based on structural MRI: The CADDementia challenge. In : *NeuroImage* 111 (2015), pp. 562–579 [16](#), [32](#), [107](#)

- [Buades et al. 2005] BUADES, Antoni ; COLL, Bartomeu ; MOREL, J-M : A non-local algorithm for image denoising. In : *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on* 2 IEEE (Veranst.), 2005, pp. 60–65 16, 19
- [Calhoun and Sui 2016] CALHOUN, Vince D. ; SUI, Jing : Multimodal fusion of brain imaging data: a key to finding the missing link (s) in complex mental illness. In : *Biological psychiatry: cognitive neuroscience and neuroimaging* 1 (2016), n. 3, pp. 230–244 59
- [Carlesimo et al. 2015] CARLESIMO, Giovanni A. ; PIRAS, Fabrizio ; ORFEI, Maria D. ; IORIO, Mariangela ; CALTAGIRONE, Carlo ; SPALLETTA, Gianfranco : Atrophy of presubiculum and subiculum is the earliest hippocampal anatomical marker of Alzheimer’s disease. In : *Alzheimer’s & Dementia: Diagnosis, Assessment & Disease Monitoring* 1 (2015), n. 1, pp. 24–32 97
- [Cawley and Talbot 2004] CAWLEY, Gavin C. ; TALBOT, Nicola L. : Fast exact leave-one-out cross-validation of sparse least-squares support vector machines. In : *Neural networks* 17 (2004), n. 10, pp. 1467–1475 23
- [Chincarini et al. 2011] CHINCARINI, Andrea ; BOSCO, Paolo ; CALVINI, Piero ; GEMME, Gianluca ; ESPOSITO, Mario ; OLIVIERI, Chiara ; REI, Luca ; SQUARCIA, Sandro ; RODRIGUEZ, Guido ; BELLOTTI, Roberto et al. : Local MRI analysis approach in the diagnosis of early and prodromal Alzheimer’s disease. In : *Neuroimage* 58 (2011), n. 2, pp. 469–480 64
- [Chu et al. 2012] CHU, Carlton ; HSU, Ai-Ling ; CHOU, Kun-Hsien ; BANDETTINI, Peter ; LIN, ChingPo ; INITIATIVE, Alzheimer’s Disease N. et al. : Does feature selection improve classification accuracy? Impact of sample size and feature selection on classification using anatomical magnetic resonance images. In : *Neuroimage* 60 (2012), n. 1, pp. 59–70 21
- [Collins et al. 1995] COLLINS, D L. ; HOLMES, Colin J. ; PETERS, Terrence M. ; EVANS, Alan C. : Automatic 3-D model-based neuroanatomical segmentation. In : *Human brain mapping* 3 (1995), n. 3, pp. 190–208 19
- [Collins and Pruessner 2010] COLLINS, D L. ; PRUESSNER, Jens C. : Towards accurate, automatic segmentation of the hippocampus and amygdala from MRI by augmenting ANIMAL with a template library and label fusion. In : *Neuroimage* 52 (2010), n. 4, pp. 1355–1366 19
- [Collins et al. 1994] COLLINS, DL ; NEELIN, P ; PETERS, TM ; EVANS, AC : *Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space.* 1994 18

- [Colliot et al. 2008] COLLIOT, Olivier ; CHÉTELAT, Gaël ; CHUPIN, Marie ; DESGRANGES, Béatrice ; MAGNIN, Benoît ; BENALI, Habib ; DUBOIS, Bruno ; GARNERO, Line ; EUSTACHE, Francis ; LEHÉRICY, Stéphane : Discrimination between Alzheimer disease, mild cognitive impairment, and normal aging by using automated segmentation of the hippocampus. In : *Radiology* 248 (2008), n. 1, pp. 194–201 [35](#), [142](#)
- [Cortes and Vapnik 1995] CORTES, Corinna ; VAPNIK, Vladimir : Support-vector networks. In : *Machine learning* 20 (1995), n. 3, pp. 273–297 [21](#), [148](#)
- [Coupé et al. 2012a] COUPÉ, Pierrick ; ESKILDSEN, Simon F. ; MANJÓN, José V ; FONOV, Vladimir S. ; COLLINS, D L. ; NEUROIMAGING INITIATIVE, Alzheimer’s disease et al. : Simultaneous segmentation and grading of anatomical structures for patient’s classification: application to Alzheimer’s disease. In : *NeuroImage* 59 (2012), n. 4, pp. 3736–3747 [37](#)
- [Coupé et al. 2012b] COUPÉ, Pierrick ; ESKILDSEN, Simon F. ; MANJÓN, José V ; FONOV, Vladimir S. ; PRUESSNER, Jens C. ; ALLARD, Michèle ; COLLINS, D L. ; DISEASE NEUROIMAGING INITIATIVE Alzheimer’s : Scoring by nonlocal image patch estimator for early detection of Alzheimer’s disease. In : *NeuroImage: clinical* 1 (2012), n. 1, pp. 141–152 [ix](#), [xiii](#), [20](#), [26](#), [27](#), [30](#), [31](#), [32](#), [35](#), [37](#), [38](#), [40](#), [64](#), [67](#), [73](#), [74](#), [75](#), [78](#), [107](#), [108](#), [117](#), [118](#), [121](#), [124](#), [125](#), [134](#), [135](#), [142](#)
- [Coupé et al. 2015] COUPÉ, Pierrick ; FONOV, Vladimir S. ; BERNARD, Charlotte ; ZANDIFAR, Azar ; ESKILDSEN, Simon F. ; HELMER, Catherine ; MANJÓN, José V ; AMIEVA, Hélène ; DARTIGUES, Jean-François ; ALLARD, Michèle : Detection of Alzheimer’s disease signature in MR images seven years before conversion to dementia: Toward an early individual prognosis. In : *Human brain mapping* 36 (2015), n. 12, pp. 4758–4770 [23](#)
- [Coupé et al. 2013] COUPÉ, Pierrick ; MANJÓN, José V ; CHAMBERLAND, Maxime ; DESCOTEAUX, Maxime ; HIBA, Bassem : Collaborative patch-based super-resolution for diffusion-weighted images. In : *NeuroImage* 83 (2013), pp. 245–261 [90](#)
- [Coupé et al. 2011] COUPÉ, Pierrick ; MANJÓN, José V ; FONOV, Vladimir ; PRUESSNER, Jens ; ROBLES, Montserrat ; COLLINS, D L. : Patch-based segmentation using expert priors: Application to hippocampus and ventricle segmentation. In : *NeuroImage* 54 (2011), n. 2, pp. 940–954 [19](#), [40](#), [57](#), [66](#)
- [Coupé et al. 2008] COUPÉ, Pierrick ; YGER, Pierre ; PRIMA, Sylvain ; HELLIER, Pierre ; KERVRANN, Charles ; BARILLOT, Christian : An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images. In : *IEEE transactions on medical imaging* 27 (2008), n. 4, pp. 425–441 [17](#), [40](#)

- [Cui et al. 2012] CUI, Yue ; WEN, Wei ; LIPNICKI, Darren M. ; BEG, Mirza F. ; JIN, Jesse S. ; LUO, Suhuai ; ZHU, Wanlin ; KOCHAN, Nicole A. ; REPPERMUND, Simone ; ZHUANG, Lin : Automated detection of amnesic mild cognitive impairment in community-dwelling elderly adults: a combined spatial atrophy and white matter alteration approach. In : *Neuroimage* 59 (2012), n. 2, pp. 1209–1217 [78](#)
- [Cuingnet et al. 2011] CUINGNET, Rémi ; GERARDIN, Emilie ; TESSIERAS, Jérôme ; AUZIAS, Guillaume ; LEHÉRICY, Stéphane ; HABERT, Marie-Odile ; CHUPIN, Marie ; BENALI, Habib ; COLLIOT, Olivier ; INITIATIVE, Alzheimer’s Disease N. : Automatic classification of patients with Alzheimer’s disease from structural MRI: a comparison of ten methods using the ADNI database. In : *Neuroimage* 56 (2011), n. 2, pp. 766–781 [16](#), [23](#), [27](#), [35](#), [142](#)
- [Cuingnet et al. 2013] CUINGNET, Rémi ; GLAUNÈS, Joan A. ; CHUPIN, Marie ; BENALI, Habib ; COLLIOT, Olivier : Spatial and anatomical regularization of SVM: a general framework for neuroimaging data. In : *IEEE transactions on pattern analysis and machine intelligence* 35 (2013), n. 3, pp. 682–696 [30](#), [35](#), [108](#), [142](#)
- [Davatzikos 1997] DAVATZIKOS, Christos : Spatial transformation and registration of brain images using elastically deformable models. In : *Computer Vision and Image Understanding* 66 (1997), n. 2, pp. 207–222 [18](#)
- [Davies et al. 2002] DAVIES, Rhodri H. ; TWINING, Carole J. ; COOTES, Timothy F. ; WATERTON, John C. ; TAYLOR, Christopher J. : A minimum description length approach to statistical shape modeling. In : *IEEE transactions on medical imaging* 21 (2002), n. 5, pp. 525–537 [27](#)
- [De Lathauwer et al. 1994] DE LATHAUWER, L ; DE MOOR, B ; VANDEWALLE, J ; HIGHER-ORDER, Blind Source S. by : Singular value decomposition. In : *Proc. EUSIPCO-94, Edinburgh, Scotland, UK* 1, 1994, pp. 175–178 [20](#)
- [DeCarli 2003] DECARLI, Charles : Mild cognitive impairment: prevalence, prognosis, aetiology, and treatment. In : *The Lancet Neurology* 2 (2003), n. 1, pp. 15–21 [3](#)
- [Dickerson et al. 2001] DICKERSON, Bradford C. ; GONCHAROVA, I ; SULLIVAN, MP ; FORCHETTI, C ; WILSON, RS ; BENNETT, DA ; BECKETT, Laurel A. et al. : MRI-derived entorhinal and hippocampal atrophy in incipient and very mild Alzheimer’s disease. In : *Neurobiology of aging* 22 (2001), n. 5, pp. 747–754 [26](#)
- [Donoho 1995] DONOHO, David L. : De-noising by soft-thresholding. In : *IEEE transactions on information theory* 41 (1995), n. 3, pp. 613–627 [16](#)

- [Donoho and Johnstone 1994] DONOHO, David L. ; JOHNSTONE, Jain M. : Ideal spatial adaptation by wavelet shrinkage. In : *biometrika* 81 (1994), n. 3, pp. 425–455 [16](#)
- [Dubois and Albert 2004] DUBOIS, Bruno ; ALBERT, Martin L. : Amnestic MCI or prodromal Alzheimer’s disease? In : *The Lancet Neurology* 3 (2004), n. 4, pp. 246–248 [4](#)
- [Dukart et al. 2011] DUKART, Juergen ; SCHROETER, Matthias L. ; MUELLER, Karsten ; DISEASE NEUROIMAGING INITIATIVE Alzheimer’s : Age correction in dementia—matching to a healthy brain. In : *PloS one* 6 (2011), n. 7, pp. e22193 [68](#), [81](#), [92](#)
- [Dyrba et al. 2015] DYRBA, Martin ; BARKHOF, Frederik ; FELLGIEBEL, Andreas ; FILIPPI, Massimo ; HAUSNER, Lucrezia ; HAUENSTEIN, Karlheinz ; KIRSTE, Thomas ; TEIPEL, Stefan J. : Predicting Prodromal Alzheimer’s Disease in Subjects with Mild Cognitive Impairment Using Machine Learning Classification of Multimodal Multicenter Diffusion-Tensor and Magnetic Resonance Imaging Data. In : *Journal of Neuroimaging* 25 (2015), n. 5, pp. 738–747 [59](#), [78](#)
- [Dyrba et al. 2013] DYRBA, Martin ; EWERS, Michael ; WEGRZYN, Martin ; KILIMANN, Ingo ; PLANT, Claudia ; OSWALD, Annahita ; MEINDL, Thomas ; PIEVANI, Michela ; BOKDE, Arun L. ; FELLGIEBEL, Andreas : Robust automated detection of microstructural white matter degeneration in Alzheimer’s disease using machine learning classification of multicenter DTI data. In : *PloS one* 8 (2013), n. 5, pp. e64925 [78](#), [98](#)
- [Dyrby et al. 2014] DYRBY, Tim B. ; LUNDELL, Henrik ; BURKE, Mark W. ; REISLEV, Nina L. ; PAULSON, Olaf B. ; PTITO, Maurice ; SIEBNER, Hartwig R. : Interpolation of diffusion weighted imaging datasets. In : *NeuroImage* 103 (2014), pp. 202–213 [81](#)
- [Ellis et al. 2009] ELLIS, Kathryn A. ; BUSH, Ashley I. ; DARBY, David ; DE FAZIO, Daniela ; FOSTER, Jonathan ; HUDSON, Peter ; LAUTENSCHLAGER, Nicola T. ; LENZO, Nat ; MARTINS, Ralph N. ; MARUFF, Paul et al. : The Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging: methodology and baseline characteristics of 1112 individuals recruited for a longitudinal study of Alzheimer’s disease. In : *International Psychogeriatrics* 21 (2009), n. 4, pp. 672–687 [22](#)
- [Eskildsen et al. 2013] ESKILDSEN, Simon F. ; COUPÉ, Pierrick ; GARCÍA-LORENZO, Daniel ; FONOV, Vladimir ; PRUESSNER, Jens C. ; COLLINS, D L. ; DISEASE NEUROIMAGING INITIATIVE Alzheimer’s : Prediction of Alzheimer’s disease in subjects with mild cognitive impairment from the ADNI cohort using patterns of cortical thinning. In : *Neuroimage* 65 (2013), pp. 511–521 [28](#), [35](#)

- [Eskildsen and Østergaard 2006] ESKILDSEN, Simon F. ; ØSTERGAARD, Lasse R. : Active surface approach for extraction of the human cerebral cortex from MRI. In : *International Conference on Medical Image Computing and Computer-Assisted Intervention* Springer (Veranst.), 2006, pp. 823–830 28
- [Fazlollahi et al. 2014] FAZLOLLAHI, Amir ; MERIAUDEAU, Fabrice ; VILLEMAGNE, Victor L. ; ROWE, Christopher C. ; YATES, Paul ; SALVADO, Olivier ; BOURGEAT, Pierrick : Efficient machine learning framework for computer-aided detection of cerebral microbleeds using the radon transform. In : *Biomedical Imaging (ISBI), 2014 IEEE 11th International Symposium on IEEE* (Veranst.), 2014, pp. 113–116 129
- [Fellgiebel et al. 2006] FELLGIEBEL, Andreas ; DELLANI, Paulo R. ; GREVERUS, Dirk ; SCHEURICH, Armin ; STOETER, Peter ; MÜLLER, Matthias J. : Predicting conversion to dementia in mild cognitive impairment by volumetric and diffusivity measurements of the hippocampus. In : *Psychiatry Research: Neuroimaging* 146 (2006), n. 3, pp. 283–287 78
- [Fellgiebel and Yakushev 2011] FELLGIEBEL, Andreas ; YAKUSHEV, Igor : Diffusion tensor imaging of the hippocampus in MCI and early Alzheimer’s disease. In : *Journal of Alzheimer’s Disease* 26 (2011), n. s3, pp. 257–262 78
- [Fischl 2012] FISCHL, Bruce : FreeSurfer. In : *Neuroimage* 62 (2012), n. 2, pp. 774–781 26
- [Fischl and Dale 2000] FISCHL, Bruce ; DALE, Anders M. : Measuring the thickness of the human cerebral cortex from magnetic resonance images. In : *Proceedings of the National Academy of Sciences* 97 (2000), n. 20, pp. 11050–11055 27, 28
- [Fisher 1936] FISHER, Ronald A. : The use of multiple measurements in taxonomic problems. In : *Annals of eugenics* 7 (1936), n. 2, pp. 179–188 21, 148
- [Folstein et al. 1975] FOLSTEIN, Marshal F. ; FOLSTEIN, Susan E. ; MCHUGH, Paul R. : “Mini-mental state”: a practical method for grading the cognitive state of patients for the clinician. In : *Journal of psychiatric research* 12 (1975), n. 3, pp. 189–198 5
- [Frisoni et al. 2010] FRISONI, Giovanni B. ; FOX, Nick C. ; JACK, Clifford R. ; SCHELTENS, Philip ; THOMPSON, Paul M. : The clinical use of structural MRI in Alzheimer disease. In : *Nature Reviews Neurology* 6 (2010), n. 2, pp. 67–77 3, 26
- [Garyfallidis et al. 2014] GARYFALLIDIS, Eleftherios ; BRETT, Matthew ; AMIRBEKIAN, Bagrat ; ROKEM, Ariel ; VAN DER WALT, Stefan ; DESCOTEAUX, Maxime ; NIMMO-SMITH, Ian : Dipy, a library for the analysis of diffusion MRI data. In : *Frontiers in neuroinformatics* 8 (2014), pp. 8 80, 91

- [Gauriau et al. 2015] GAURIAU, Romane ; CUINGNET, Rémi ; LESAGE, David ; BLOCH, Isabelle : Multi-organ localization with cascaded global-to-local regression and shape prior. In : *Medical image analysis* 23 (2015), n. 1, pp. 70–83 [129](#)
- [Geisser 1975] GEISSER, Seymour : The predictive sample reuse method with applications. In : *Journal of the American statistical Association* 70 (1975), n. 350, pp. 320–328 [23](#)
- [Geman and Geman 1984] GEMAN, Stuart ; GEMAN, Donald : Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. In : *IEEE Transactions on pattern analysis and machine intelligence* (1984), n. 6, pp. 721–741 [16](#)
- [Gerardin et al. 2009] GERARDIN, Emilie ; CHÉTELAT, Gaël ; CHUPIN, Marie ; CUINGNET, Rémi ; DESGRANGES, Béatrice ; KIM, Ho-Sung ; NIETHAMMER, Marc ; DUBOIS, Bruno ; LEHÉRICY, Stéphane ; GARNERO, Line : Multidimensional classification of hippocampal shape features discriminates Alzheimer’s disease and mild cognitive impairment from normal aging. In : *Neuroimage* 47 (2009), n. 4, pp. 1476–1486 [27](#), [35](#), [142](#)
- [Giraud et al. 2016] GIRAUD, Rémi ; TA, Vinh-Thong ; PAPADAKIS, Nicolas ; MANJÓN, José V ; COLLINS, D L. ; COUPÉ, Pierrick ; DISEASE NEUROIMAGING INITIATIVE Alzheimer’s : An optimized patchmatch for multi-scale and multi-feature label fusion. In : *NeuroImage* 124 (2016), pp. 770–782 [20](#), [42](#), [43](#), [57](#), [59](#), [60](#), [62](#), [64](#), [66](#), [81](#), [92](#), [114](#), [122](#), [123](#), [130](#)
- [Gray et al. 2013] GRAY, Katherine R. ; ALJABAR, Paul ; HECKEMANN, Rolf A. ; HAMMERS, Alexander ; RUECKERT, Daniel ; INITIATIVE, Alzheimer’s Disease N. et al. : Random forest-based similarity measures for multi-modal classification of Alzheimer’s disease. In : *NeuroImage* 65 (2013), pp. 167–175 [59](#), [129](#)
- [Grigorescu et al. 2002] GRIGORESCU, Simona E. ; PETKOV, Nicolai ; KRUIZINGA, Peter : Comparison of texture features based on Gabor filters. In : *IEEE Transactions on Image processing* 11 (2002), n. 10, pp. 1160–1167 [64](#)
- [Hardy 2006] HARDY, John : Alzheimer’s disease: the amyloid cascade hypothesis: an update and reappraisal. In : *Journal of Alzheimer’s disease* 9 (2006), n. s3, pp. 151–153 [3](#)
- [Heckemann et al. 2006] HECKEMANN, Rolf A. ; HAJNAL, Joseph V. ; ALJABAR, Paul ; RUECKERT, Daniel ; HAMMERS, Alexander : Automatic anatomical brain MRI segmentation combining label propagation and decision fusion. In : *NeuroImage* 33 (2006), n. 1, pp. 115–126 [19](#)

- [Hett et al. 2016] HETT, Kilian ; TA, Vinh-Thong ; GIRAUD, Rémi ; MONDINO, Mary ; MANJÓN, José V ; COUPÉ, Pierrick ; DISEASE NEUROIMAGING INITIATIVE Alzheimer's : Patch-based DTI grading: Application to Alzheimer's disease classification. In : *International Workshop on Patch-based Techniques in Medical Imaging* Springer (Veranst.), 2016, pp. 76–83 42
- [Hett et al. 2018a] HETT, Kilian ; TA, Vinh-Thong ; MANJÓN, José V ; COUPÉ, Pierrick : Graph of hippocampal subfields grading for Alzheimer's disease prediction. In : *International Workshop on Machine Learning in Medical Imaging* Springer (Veranst.), 2018, pp. 259–266 xiii, 142
- [Hett et al. 2017] HETT, Kilian ; TA, Vinh-Thong ; MANJÓN, José V ; COUPÉ, Pierrick ; INITIATIVE, Alzheimer's Disease N. : Adaptive fusion of texture-based grading: Application to Alzheimer's disease detection. In : *International Workshop on Patch-based Techniques in Medical Imaging* Springer (Veranst.), 2017, pp. 82–89 108
- [Hett et al. 2018b] HETT, Kilian ; TA, Vinh-Thong ; MANJÓN, José V ; COUPÉ, Pierrick ; INITIATIVE, Alzheimer's Disease N. et al. : Graph of brain structures grading for early detection of Alzheimer's disease. In : *International Conference on Medical Image Computing and Computer-Assisted Intervention* Springer (Veranst.), 2018, pp. 429–436 xiii, 142
- [Hett et al. 2018c] HETT, Kilian ; TA, Vinh-Thong ; MANJÓN, José V ; COUPÉ, Pierrick ; INITIATIVE, Alzheimer's Disease N. et al. : Multiple anatomical graph of structure grading for Alzheimer's disease prediction. In : *NeuroImage (Submitted)* (2018) xiii, 142
- [Hett et al. 2018d] HETT, Kilian ; VINH-THONG, TA ; MANJÓN, José V ; COUPÉ, Pierrick : Adaptive fusion of texture-based grading for Alzheimer's disease classification. In : *Computerized Medical Imaging and Graphics* (2018) xiii, 142
- [Hinrichs et al. 2011] HINRICHS, Chris ; SINGH, Vikas ; XU, Guofan ; JOHNSON, Sterling C. ; INITIATIVE, Alzheimers Disease N. et al. : Predictive markers for AD in a multi-modality framework: an analysis of MCI progression in the ADNI population. In : *Neuroimage* 55 (2011), n. 2, pp. 574–589 59
- [Hirata et al. 2005] HIRATA, Yoko ; MATSUDA, Hiroshi ; NEMOTO, Kiyotaka ; OHNISHI, Takashi ; HIRAO, Kentaro ; YAMASHITA, Fumio ; ASADA, Takashi ; IWABUCHI, Satoshi ; SAMEJIMA, Hirotsugu : Voxel-based morphometry to discriminate early Alzheimer's disease from controls. In : *Neuroscience letters* 382 (2005), n. 3, pp. 269–274 30
- [Hochberg and Tamhane 1987] HOCHBERG, Yosef ; TAMHANE, Ajit : *Multiple comparison procedures*. John Wiley, 1987 94

- [Hua et al. 2008] HUA, Xue ; LEOW, Alex D. ; PARIKSHAK, Neelroop ; LEE, Suh ; CHIANG, Ming-Chang ; TOGA, Arthur W. ; JACK JR, Clifford R. ; WEINER, Michael W. ; THOMPSON, Paul M. ; INITIATIVE, Alzheimer’s Disease N. et al. : Tensor-based morphometry as a neuroimaging biomarker for Alzheimer’s disease: an MRI study of 676 AD, MCI, and normal subjects. In : *Neuroimage* 43 (2008), n. 3, pp. 458–469 30
- [Hutton et al. 2008] HUTTON, Chloe ; DE VITA, Enrico ; ASHBURNER, John ; DEICHMANN, Ralf ; TURNER, Robert : Voxel-based cortical thickness measurements in MRI. In : *Neuroimage* 40 (2008), n. 4, pp. 1701–1710 27
- [Iglesias and Sabuncu 2015] IGLESIAS, Juan E. ; SABUNCU, Mert R. : Multi-atlas segmentation of biomedical images: a survey. In : *Medical image analysis* 24 (2015), n. 1, pp. 205–219 19
- [Jack et al. 2008] JACK, Clifford R. ; BERNSTEIN, Matt A. ; FOX, Nick C. ; THOMPSON, Paul ; ALEXANDER, Gene ; HARVEY, Danielle ; BOROWSKI, Bret ; BRITSON, Paula J. ; L WHITWELL, Jennifer ; WARD, Chadwick : The Alzheimer’s disease neuroimaging initiative (ADNI): MRI methods. In : *Journal of magnetic resonance imaging* 27 (2008), n. 4, pp. 685–691 22, 79
- [Jack et al. 1992] JACK, Clifford R. ; PETERSEN, Ronald C. ; O’BRIEN, Peter C. ; TANGALOS, Eric G. : MR-based hippocampal volumetry in the diagnosis of Alzheimer’s disease. In : *Neurology* 42 (1992), n. 1, pp. 183–183 3, 26
- [Jack Jr et al. 2011] JACK JR, Clifford R. ; BARKHOF, Frederik ; BERNSTEIN, Matt A. ; CANTILLON, Marc ; COLE, Patricia E. ; DECARLI, Charles ; DUBOIS, Bruno ; DUCHESNE, Simon ; FOX, Nick C. ; FRISONI, Giovanni B. et al. : Steps to standardization and validation of hippocampal volumetry as a biomarker in clinical trials and diagnostic criterion for Alzheimer’s disease. In : *Alzheimer’s & Dementia* 7 (2011), n. 4, pp. 474–485 26
- [Jahanshad et al. 2010] JAHANSHAD, Neda ; ZHAN, Liang ; BERNSTEIN, Matt A. ; BOROWSKI, Bret J. ; JACK, Clifford R. ; TOGA, Arthur W. ; THOMPSON, Paul M. : Diffusion tensor imaging in seven minutes: determining trade-offs between spatial and directional resolution. In : *Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on IEEE* (Veranst.), 2010, pp. 1161–1164 79
- [Janoušová et al. 2012] JANOUŠOVÁ, Eva ; VOUNOU, Maria ; WOLZ, Robin ; GRAY, Katherine R. ; RUECKERT, Daniel ; MONTANA, Giovanni et al. : Biomarker discovery for sparse classification of brain images in Alzheimer’s disease. In : *Annals of the BMVA* (2012), n. 2 38

BIBLIOGRAPHY

- [Jenkinson et al. 2012] JENKINSON, Mark ; BECKMANN, Christian F. ; BEHRENS, Timothy E. ; WOOLRICH, Mark W. ; SMITH, Stephen M. : Fsl. In : *Neuroimage* 62 (2012), n. 2, pp. 782–790 26
- [Jolliffe 2011] JOLLIFFE, Ian : Principal component analysis. In : *International encyclopedia of statistical science*. Springer, 2011, pp. 1094–1096 20
- [Kantarci et al. 2005] KANTARCI, Kejal ; PETERSEN, Ronald C. ; BOEVE, Bradley F. ; KNOPMAN, David S. ; WEIGAND, Stephen D. ; O'BRIEN, Peter C. ; SHIUNG, Maria M. ; SMITH, Glenn E. ; IVNIK, Robert J. ; TANGALOS, Eric G. : DWI predicts future progression to Alzheimer disease in amnesic mild cognitive impairment. In : *Neurology* 64 (2005), n. 5, pp. 902–904 78
- [Karas et al. 2003] KARAS, GB ; BURTON, EJ ; ROMBOUTS, SARB ; VAN SCHIJNDEL, RA ; O'BRIEN, JTh ; SCHELTENS, PH ; MCKEITH, IG ; WILLIAMS, D ; BALLARD, C ; BARKHOF, F : A comprehensive study of gray matter loss in patients with Alzheimer's disease using optimized voxel-based morphometry. In : *Neuroimage* 18 (2003), n. 4, pp. 895–907 30
- [Kerchner et al. 2010] KERCHNER, GA ; HESS, CP ; HAMMOND-ROSENBLUTH, KE ; XU, D ; RABINOVICI, GD ; KELLEY, DAC ; VIGNERON, DB ; NELSON, SJ ; MILLER, BL : Hippocampal CA1 apical neuropil atrophy in mild Alzheimer disease visualized with 7-T MRI. In : *Neurology* 75 (2010), n. 15, pp. 1381–1387 89, 97
- [Kerchner et al. 2012] KERCHNER, Geoffrey A. ; DEUTSCH, Gayle K. ; ZEINEH, Michael ; DOUGHERTY, Robert F. ; SARANATHAN, Manojkumar ; RUTT, Brian K. : Hippocampal CA1 apical neuropil atrophy and memory performance in Alzheimer's disease. In : *Neuroimage* 63 (2012), n. 1, pp. 194–202 89, 97, 116
- [Kesslak et al. 1991] KESSLAK, J P. ; NALCIOGLU, Orhan ; COTMAN, Carl W. : Quantification of magnetic resonance scans for hippocampal and parahippocampal atrophy in Alzheimer's disease. In : *Neurology* 41 (1991), n. 1, pp. 51–51 26
- [Khan et al. 2015] KHAN, Wasim ; WESTMAN, Eric ; JONES, Nigel ; WAHLUND, Lars-Olof ; MECOCCI, Patrizia ; VELLAS, Bruno ; TSOLAKI, Magda ; KŁOSZEWSKA, Iwona ; SOININEN, Hilikka ; SPENGER, Christian et al. : Automated hippocampal subfield measures as predictors of conversion from mild cognitive impairment to Alzheimer's disease in two independent cohorts. In : *Brain topography* 28 (2015), n. 5, pp. 746–759 94
- [Kim et al. 2013] KIM, Minjeong ; WU, Guorong ; LI, Wei ; WANG, Li ; SON, Young-Don ; CHO, Zang-Hee ; SHEN, Dinggang : Automatic hippocampus segmentation of 7.0 Tesla MR images by combining multiple atlases and auto-context models. In : *NeuroImage* 83 (2013), pp. 335–345 59

- [Kohannim et al. 2010] KOHANNIM, Omid ; HUA, Xue ; HIBAR, Derrek P. ; LEE, Suh ; CHOU, Yi-Yu ; TOGA, Arthur W. ; JACK JR, Clifford R. ; WEINER, Michael W. ; THOMPSON, Paul M. ; INITIATIVE, Alzheimer's Disease N. et al. : Boosting power for clinical trials using classifiers based on multiple biomarkers. In : *Neurobiology of aging* 31 (2010), n. 8, pp. 1429–1442 59
- [Koikkalainen et al. 2012] KOIKKALAINEN, Juha ; PÖLÖNEN, Harri ; MATTILA, Jussi ; VAN GILS, Mark ; SOININEN, Hilka ; LÖTJÖNEN, Jyrki ; INITIATIVE, Alzheimer's Disease N. et al. : Improved classification of Alzheimer's disease data via removal of nuisance variability. In : *PloS one* 7 (2012), n. 2, pp. e31112 108
- [Komlagan et al. 2014] KOMLAGAN, Mawulawoé ; TA, Vinh-Thong ; PAN, Xingyu ; DOMENGER, Jean-Philippe ; COLLINS, D L. ; COUPÉ, Pierrick ; DISEASE NEUROIMAGING INITIATIVE Alzheimer's : Anatomically constrained weak classifier fusion for early detection of Alzheimer's disease. In : *International Workshop on Machine Learning in Medical Imaging* Springer (Veranst.), 2014, pp. 141–148 64, 78
- [Kuroda et al. 2011] KURODA, Masahiro ; MORI, Yuichi ; IIZUKA, Masaya ; SAKAKIHARA, Michio : Acceleration of the alternating least squares algorithm for principal components analysis. In : *Computational Statistics & Data Analysis* 55 (2011), n. 1, pp. 143–153 20
- [La Joie et al. 2013] LA JOIE, Renaud ; PERROTIN, Audrey ; DE LA SAYETTE, Vincent ; EGRET, Stéphanie ; DOEUVRE, Loïc ; BELLIARD, Serge ; EUSTACHE, Francis ; DESGRANGES, Béatrice ; CHÉTELAT, Gaël : Hippocampal subfield volumetry in mild cognitive impairment, Alzheimer's disease and semantic dementia. In : *NeuroImage: Clinical* 3 (2013), pp. 155–162 89, 97
- [La Rocca et al. 2018] LA ROCCA, Marianna ; AMOROSO, Nicola ; MONACO, Alfonso ; BELLOTTI, Roberto ; TANGARO, Sabina : A novel approach to brain connectivity reveals early structural changes in Alzheimer's disease. In : *Physiological Measurement* (2018) 85, 96
- [Lachenbruch and Goldstein 1979] LACHENBRUCH, Peter A. ; GOLDSTEIN, M : Discriminant analysis. In : *Biometrics* (1979), pp. 69–85 21, 148
- [Larson 1931] LARSON, Selmer C. : The shrinkage of the coefficient of multiple correlation. In : *Journal of Educational Psychology* 22 (1931), n. 1, pp. 45 23
- [LeCun et al. 2015] LECUN, Yann ; BENGIO, Yoshua ; HINTON, Geoffrey : Deep learning. In : *nature* 521 (2015), n. 7553, pp. 436 21, 151
- [Ledig et al. 2018] LEDIG, Christian ; SCHUH, Andreas ; GUERRERO, Ricardo ; HECKEMANN, Rolf A. ; RUECKERT, Daniel : Structural brain imaging in Alzheimer's disease and mild cognitive impairment: biomarker analysis and shared morphometry database. In : *Scientific reports* 8 (2018), n. 1, pp. 11258 26

BIBLIOGRAPHY

- [de Leon et al. 2001] LEON, Mony de ; BOBINSKI, Maciek ; CONVIT, Antonio ; WOLF, Oliver ; INSAUSTI, Ricardo : Usefulness of MRI measures of entorhinal cortex versus hippocampus in AD. In : *Neurology* 56 (2001), n. 6, pp. 820–823 [26](#)
- [Li et al. 2018] LI, Fan ; LIU, Manhua ; INITIATIVE, Alzheimer’s Disease N. et al. : Alzheimer’s Disease Diagnosis Based on Multiple Cluster Dense Convolutional Networks. In : *Computerized Medical Imaging and Graphics* (2018) [23](#), [32](#), [33](#), [35](#), [142](#)
- [Li et al. 2014] LI, Muwei ; QIN, Yuanyuan ; GAO, Fei ; ZHU, Wenzhen ; HE, Xiaohai : Discriminative analysis of multivariate features from structural MRI and diffusion tensor images. In : *Magnetic resonance imaging* 32 (2014), n. 8, pp. 1043–1051 [78](#), [98](#)
- [Li et al. 2013] LI, Ya-Di ; DONG, Hai-Bo ; XIE, Guo-Ming ; ZHANG, Ling-jun : Discriminative analysis of mild Alzheimer’s disease and normal aging using volume of hippocampal subfields and hippocampal mean diffusivity: an in vivo magnetic resonance imaging study. In : *American Journal of Alzheimer’s Disease & Other Dementias* 28 (2013), n. 6, pp. 627–633 [89](#), [97](#)
- [Liang and Lauterbur 2000] LIANG, Zhi-Pei ; LAUTERBUR, Paul C. : *Principles of magnetic resonance imaging: a signal processing perspective*. SPIE Optical Engineering Press, 2000 [17](#)
- [Liao and Chung 2010] LIAO, Shu ; CHUNG, Albert C. : Feature based nonrigid brain MR image registration with symmetric alpha stable filters. In : *IEEE transactions on medical imaging* 29 (2010), n. 1, pp. 106–119 [18](#)
- [Liu et al. 2009] LIU, Jun ; JI, Shuiwang ; YE, Jieping et al. : SLEP: Sparse learning with efficient projections. In : *Arizona State University* 6 (2009), n. 491, pp. 7 [115](#), [123](#), [131](#)
- [Liu et al. 2012] LIU, Manhua ; ZHANG, Daoqiang ; SHEN, Dinggang ; DISEASE NEUROIMAGING INITIATIVE Alzheimer’s : Ensemble sparse classification of Alzheimer’s disease. In : *NeuroImage* 60 (2012), n. 2, pp. 1106–1116 [xiii](#), [23](#), [57](#), [64](#), [73](#), [74](#), [75](#), [78](#), [134](#), [135](#)
- [Liu et al. 2014] LIU, Manhua ; ZHANG, Daoqiang ; SHEN, Dinggang ; INITIATIVE, Alzheimer’s Disease N. : Hierarchical fusion of features and classifier decisions for Alzheimer’s disease diagnosis. In : *Human brain mapping* 35 (2014), n. 4, pp. 1305–1319 [108](#)
- [Liu et al. 2018] LIU, Mingxia ; ZHANG, Jun ; ADELI, Ehsan ; SHEN, Dinggang : Landmark-based deep multi-instance learning for brain disease diagnosis. In : *Medical image analysis* 43 (2018), pp. 157–168 [32](#), [33](#), [35](#), [142](#)

- [Liu et al. 2011] LIU, Yawu ; SPULBER, Gabriela ; LEHTIMÄKI, Kimmo K. ; KÖNÖNEN, Mervi ; HALLIKAINEN, Ilona ; GRÖHN, Heidi ; KIVIPELTO, Miia ; HALLIKAINEN, Merja ; VANNINEN, Ritva ; SOININEN, Hilikka : Diffusion tensor imaging and tract-based spatial statistics in Alzheimer’s disease and mild cognitive impairment. In : *Neurobiology of aging* 32 (2011), n. 9, pp. 1558–1571 [78](#)
- [Luo et al. 2017] LUO, Suhuai ; LI, Xuechen ; LI, Jiaming : Automatic Alzheimer’s Disease Recognition from MRI Data Using Deep Learning Method. In : *Journal of Applied Mathematics and Physics* 5 (2017), n. 09, pp. 1892 [32](#), [33](#), [35](#), [142](#)
- [Maggipinto et al. 2017] MAGGIPINTO, Tommaso ; BELLOTTI, Roberto ; AMOROSO, Nicola ; DIACONO, Domenico ; DONVITO, Giacinto ; LELLA, Eufemia ; MONACO, Alfonso ; SCELSI, Marzia A. ; TANGARO, Sabina ; INITIATIVE, Alzheimer’s Disease N. : DTI measurements for Alzheimer’s classification. In : *Physics in Medicine and Biology* 62 (2017), n. 6, pp. 2361 [84](#), [85](#), [96](#)
- [Manjón et al. 2014] MANJÓN, José ; ESKILDSEN, Simon ; COUPÉ, Pierrick ; ROMERO, José ; COLLINS, Louis ; ROBLES, Montserrat : NICE: non-local intracranial cavity extraction. In : *International Journal of Biomedical Imaging* (2014) [65](#), [90](#)
- [Manjón et al. 2008] MANJÓN, José V ; CARBONELL-CABALLERO, José ; LULL, Juan J. ; GARCÍA-MARTÍ, Gracián ; MARTÍ-BONMATÍ, Luís ; ROBLES, Montserrat : MRI denoising using non-local means. In : *Medical image analysis* 12 (2008), n. 4, pp. 514–523 [17](#)
- [Manjón and Coupé 2016] MANJÓN, José V ; COUPÉ, Pierrick : volBrain: An online MRI brain volumetry system. In : *Frontiers in neuroinformatics* 10 (2016) [26](#), [65](#), [79](#), [90](#)
- [Manjón et al. 2013] MANJÓN, José V ; COUPÉ, Pierrick ; CONCHA, Luis ; BUADES, Antonio ; COLLINS, D L. ; ROBLES, Montserrat : Diffusion weighted image denoising using overcomplete local PCA. In : *PloS one* 8 (2013), n. 9, pp. e73021 [80](#), [91](#)
- [Manjón et al. 2010] MANJÓN, José V ; COUPÉ, Pierrick ; MARTÍ-BONMATÍ, Luis ; COLLINS, D L. ; ROBLES, Montserrat : Adaptive non-local means denoising of MR images with spatially varying noise levels. In : *Journal of Magnetic Resonance Imaging* 31 (2010), n. 1, pp. 192–203 [17](#), [65](#), [79](#), [90](#), [114](#), [122](#), [130](#)
- [Manjunath and Ma 1996] MANJUNATH, Bangalore S. ; MA, Wei-Ying : Texture features for browsing and retrieval of image data. In : *IEEE Transactions on pattern analysis and machine intelligence* 18 (1996), n. 8, pp. 837–842 [64](#), [66](#)
- [Marcus et al. 2007] MARCUS, Daniel S. ; WANG, Tracy H. ; PARKER, Jamie ; CSERNANSKY, John G. ; MORRIS, John C. ; BUCKNER, Randy L. : Open

BIBLIOGRAPHY

- Access Series of Imaging Studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. In : *Journal of cognitive neuroscience* 19 (2007), n. 9, pp. 1498–1507 [22](#)
- [Marstal et al. 2016] MARSTAL, Kasper ; BERENDSEN, Floris ; STARING, Marius ; KLEIN, Stefan : SimpleElastix: A user-friendly, multi-lingual library for medical image registration. In : *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2016, pp. 134–142 [18](#)
- [Mayer et al. 2011] MAYER, Arnaldo ; ZIMMERMAN-MORENO, Gali ; SHADMI, Ran ; BATIKOFF, Amit ; GREENSPAN, Hayit : A supervised framework for the registration and segmentation of white matter fiber tracts. In : *IEEE transactions on medical imaging* 30 (2011), n. 1, pp. 131–145 [18](#)
- [McVeigh et al. 1986] MCVEIGH, ER ; BRONSKILL, MJ ; HENKELMAN, RM : Phase and sensitivity of receiver coils in magnetic resonance imaging. In : *Medical physics* 13 (1986), n. 6, pp. 806–814 [17](#)
- [Miller et al. 2003] MILLER, MI ; HOSAKERE, M ; BARKER, AR ; PRIEBE, CE ; LEE, N ; RATNANATHER, JT ; WANG, L ; GADO, M ; MORRIS, JC ; CSERNANSKY, JG : Labeled cortical mantle distance maps of the cingulate quantify differences between dementia of the Alzheimer type and healthy aging. In : *Proceedings of the National Academy of Sciences* 100 (2003), n. 25, pp. 15172–15177 [27](#)
- [Mohs 1983] MOHS, RC : The Alzheimer’s disease assessment scale: an instrument for assessing treatment efficacy. In : *Psychopharmacol Bull* 19 (1983), pp. 448–450 [5](#)
- [Moradi et al. 2015] MORADI, Elaheh ; PEPE, Antonietta ; GASER, Christian ; HUTTUNEN, Heikki ; TOHKA, Jussi ; INITIATIVE, Alzheimer’s Disease N. et al. : Machine learning framework for early MRI-based Alzheimer’s conversion prediction in MCI subjects. In : *Neuroimage* 104 (2015), pp. 398–412 [xiii](#), [30](#), [35](#), [117](#), [118](#), [121](#), [124](#), [125](#), [127](#), [134](#), [135](#), [136](#), [142](#)
- [Morris et al. 1988] MORRIS, John C. ; MCKEEL JR, Daniel W. ; FULLING, Keith ; TORACK, Richard M. ; BERG, Leonard : Validation of clinical diagnostic criteria for Alzheimer’s disease. In : *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society* 24 (1988), n. 1, pp. 17–22 [6](#)
- [Mueller et al. 2007] MUELLER, SG ; STABLES, L ; DU, AT ; SCHUFF, N ; TRURAN, D ; CASHDOLLAR, N ; WEINER, MW : Measurement of hippocampal subfields and age-related changes with high resolution MRI at 4T. In : *Neurobiology of aging* 28 (2007), n. 5, pp. 719–726 [97](#)

- [Müller et al. 2005] MÜLLER, Matthias J. ; GREVERUS, Dirk ; DELLANI, Paulo R. ; WEIBRICH, Carsten ; WILLE, Paulo R. ; SCHEURICH, Armin ; STOETER, Peter ; FELLGIEBEL, Andreas : Functional implications of hippocampal volume and diffusivity in mild cognitive impairment. In : *Neuroimage* 28 (2005), n. 4, pp. 1033–1042 78
- [Nir et al. 2013] NIR, Talia M. ; JAHANSHAD, Neda ; VILLALON-REINA, Julio E. ; TOGA, Arthur W. ; JACK, Clifford R. ; WEINER, Michael W. ; THOMPSON, Paul M. ; DISEASE NEUROIMAGING INITIATIVE (ADNI Alzheimer’s : Effectiveness of regional DTI measures in distinguishing Alzheimer’s disease, MCI, and normal aging. In : *NeuroImage: clinical* 3 (2013), pp. 180–195 78
- [Nir et al. 2015] NIR, Talia M. ; VILLALON-REINA, Julio E. ; PRASAD, Gautam ; JAHANSHAD, Neda ; JOSHI, Shantanu H. ; TOGA, Arthur W. ; BERNSTEIN, Matt A. ; JACK, Clifford R. ; WEINER, Michael W. ; THOMPSON, Paul M. : Diffusion weighted imaging-based maximum density path analysis and classification of Alzheimer’s disease. In : *Neurobiology of aging* 36 (2015), pp. S132–S140 84, 85, 96, 98
- [Lorente de Nó 1934] NÓ, Rafael Lorente de : Studies on the structure of the cerebral cortex. II. Continuation of the study of the ammonic system. In : *Journal für Psychologie und Neurologie* (1934) 89
- [O’Dwyer et al. 2012] O’DWYER, Laurence ; LAMBERTON, Franck ; BOKDE, Arun L. ; EWERS, Michael ; FALUYI, Yetunde O. ; TANNER, Colby ; MAZOYER, Bernard ; O’NEILL, Desmond ; BARTLEY, Máiréad ; COLLINS, D R. : Using support vector machines with multiple indices of diffusion for automated classification of mild cognitive impairment. In : *PloS one* 7 (2012), n. 2, pp. e32441 78
- [Oishi et al. 2011] OISHI, Kenichi ; AKHTER, Kazi ; MIELKE, Michelle ; CERITOGU, Can ; ZHANG, Jiangyang ; JIANG, Hangyi ; LI, Xin ; YOUNES, Laurent ; MILLER, Michael I. ; ZIJL, Peter van et al. : Multi-modal MRI analysis with disease-specific spatial filtering: initial testing to predict mild cognitive impairment patients who convert to Alzheimer’s disease. In : *Frontiers in neurology* 2 (2011), pp. 54 98
- [Parisot et al. 2017] PARISOT, Sarah ; KTENA, Sofia I. ; FERRANTE, Enzo ; LEE, Matthew ; MORENO, Ricardo G. ; GLOCKER, Ben ; RUECKERT, Daniel : Spectral graph convolutions for population-based disease prediction. In : *International Conference on Medical Image Computing and Computer-Assisted Intervention* Springer (Veranst.), 2017, pp. 177–185 143
- [Perona and Malik 1990] PERONA, Pietro ; MALIK, Jitendra : Scale-space and edge detection using anisotropic diffusion. In : *IEEE Transactions on pattern analysis and machine intelligence* 12 (1990), n. 7, pp. 629–639 16

BIBLIOGRAPHY

- [Petersen et al. 1999] PETERSEN, Ronald C. ; SMITH, Glenn E. ; WARING, Stephen C. ; IVNIK, Robert J. ; TANGALOS, Eric G. ; KOKMEN, Emre : Mild cognitive impairment: clinical characterization and outcome. In : *Archives of neurology* 56 (1999), n. 3, pp. 303–308 4
- [Petrella et al. 2003] PETRELLA, Jeffrey R. ; COLEMAN, R E. ; DORAISWAMY, P M. : Neuroimaging and early diagnosis of Alzheimer disease: a look to the future. In : *Radiology* 226 (2003), n. 2, pp. 315–336 2
- [Pfeffer et al. 1982] PFEFFER, Robert I. ; KUROSAKI, TT ; HARRAH JR, CH ; CHANCE, JM ; FILOS, S : Measurement of functional activities in older adults in the community. In : *Journal of gerontology* 37 (1982), n. 3, pp. 323–329 5
- [Picard and Cook 1984] PICARD, Richard R. ; COOK, R D. : Cross-validation of regression models. In : *Journal of the American Statistical Association* 79 (1984), n. 387, pp. 575–583 23
- [Pipitone et al. 2014] PIPITONE, Jon ; PARK, Min Tae M. ; WINTERBURN, Julie ; LETT, Tristram A. ; LERCH, Jason P. ; PRUESSNER, Jens C. ; LEPAGE, Martin ; VOINESKOS, Aristotle N. ; CHAKRAVARTY, M M. ; INITIATIVE, Alzheimer’s Disease N. et al. : Multi-atlas segmentation of the whole hippocampus and subfields using multiple automatically generated templates. In : *Neuroimage* 101 (2014), pp. 494–512 19
- [Portilla and Simoncelli 2003] PORTILLA, Javier ; SIMONCELLI, Eero : Image restoration using Gaussian scale mixtures in the wavelet domain. In : *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on 2 IEEE (Veranst.)*, 2003, pp. II–965 16
- [Prasad et al. 2015] PRASAD, Gautam ; JOSHI, Shantanu H. ; NIR, Talia M. ; TOGA, Arthur W. ; THOMPSON, Paul M. ; DISEASE NEUROIMAGING INITIATIVE (ADNI) Alzheimer’s : Brain connectivity and novel network measures for Alzheimer’s disease classification. In : *Neurobiology of aging* 36 (2015), pp. S121–S131 78, 84, 85, 86, 96
- [Prince et al. 2013] PRINCE, Martin ; BRYCE, Renata ; ALBANESE, Emiliano ; WIMO, Anders ; RIBEIRO, Wagner ; FERRI, Cleusa P. : The global prevalence of dementia: a systematic review and metaanalysis. In : *Alzheimer’s & dementia: the journal of the Alzheimer’s Association* 9 (2013), n. 1, pp. 63–75 2
- [Querbes et al. 2009] QUERBES, Olivier ; AUBRY, Florent ; PARIENTE, Jérémie ; LOTTERIE, Jean-Albert ; DÉMONET, Jean-François ; DURET, Véronique ; PUEL, Michèle ; BERRY, Isabelle ; FORT, Jean-Claude ; CELSIS, Pierre et al. : Early diagnosis of Alzheimer’s disease using cortical thickness: impact of cognitive reserve. In : *Brain* 132 (2009), n. 8, pp. 2036–2047 28, 35, 142

- [Quinlan 1986] QUINLAN, J. R. : Induction of decision trees. In : *Machine learning* 1 (1986), n. 1, pp. 81–106 [21](#), [150](#)
- [Raamana et al. 2015] RAAMANA, Pradeep R. ; WEINER, Michael W. ; WANG, Lei ; BEG, Mirza F. ; INITIATIVE, Alzheimer’s Disease N. et al. : Thickness network features for prognostic applications in dementia. In : *Neurobiology of aging* 36 (2015), pp. S91–S102 [28](#), [35](#), [142](#)
- [Riaz et al. 2013] RIAZ, Farhan ; HASSAN, Ali ; REHMAN, Saad ; QAMAR, Usman : Texture classification using rotation-and scale-invariant gabor texture features. In : *IEEE Signal Processing Letters* 20 (2013), n. 6, pp. 607–610 [64](#)
- [Rogers et al. 1993] ROGERS, Joseph ; KIRBY, LC ; HEMPELMAN, SR ; BERRY, DL ; MCGEER, PL ; KASZNIAK, AW ; ZALINSKI, J ; COFIELD, M ; MANSUKHANI, L ; WILLSON, P et al. : Clinical trial of indomethacin in Alzheimer’s disease. In : *Neurology* 43 (1993), n. 8, pp. 1609–1609 [5](#)
- [Rohlfing et al. 2004] ROHLFING, Torsten ; BRANDT, Robert ; MENZEL, Randolph ; MAURER JR, Calvin R. : Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains. In : *NeuroImage* 21 (2004), n. 4, pp. 1428–1442 [19](#)
- [Romero et al. 2016] ROMERO, José E ; COUPÉ, Pierrick ; MANJÓN, José V : High Resolution Hippocampus Subfield Segmentation Using Multispectral Multi-atlas Patch-Based Label Fusion. In : *International Workshop on Patch-based Techniques in Medical Imaging* Springer (Veranst.), 2016, pp. 117–124 [90](#)
- [Romero et al. 2017] ROMERO, Jose E. ; COUPE, Pierrick ; MANJON, Jose V. : HIPS: A new hippocampus subfield segmentation method. In : *NeuroImage* 163 (2017), pp. 286–295 [59](#), [60](#), [90](#), [114](#), [130](#)
- [Rose et al. 2008] ROSE, Stephen E. ; ANDREW, L ; CHALK, Jonathan B. : Gray and white matter changes in Alzheimer’s disease: a diffusion tensor imaging study. In : *Journal of Magnetic Resonance Imaging* 27 (2008), n. 1, pp. 20–26 [78](#)
- [Rubinov and Sporns 2010] RUBINOV, Mikail ; SPORNS, Olaf : Complex network measures of brain connectivity: uses and interpretations. In : *Neuroimage* 52 (2010), n. 3, pp. 1059–1069 [86](#)
- [Rubner et al. 2000] RUBNER, Yossi ; TOMASI, Carlo ; GUIBAS, Leonidas J. : The earth mover’s distance as a metric for image retrieval. In : *International journal of computer vision* 40 (2000), n. 2, pp. 99–121 [110](#)
- [Ryali et al. 2010] RYALI, Srikanth ; SUPEKAR, Kaustubh ; ABRAMS, Daniel A. ; MENON, Vinod : Sparse logistic regression for whole-brain classification of fMRI data. In : *NeuroImage* 51 (2010), n. 2, pp. 752–764 [21](#)

BIBLIOGRAPHY

- [Sabuncu et al. 2015] SABUNCU, Mert R. ; KONUKOGLU, Ender ; INITIATIVE, Alzheimer’s Disease N. et al. : Clinical prediction from structural brain MRI scans: a large-scale empirical study. In : *Neuroinformatics* 13 (2015), n. 1, pp. 31–46 16
- [Saint-Marc et al. 1989] SAINT-MARC, Philippe ; CHEN, Jer-Sen ; MEDIONI, Gérard : Adaptive smoothing: A general tool for early vision. In : *Computer Vision and Pattern Recognition, 1989. Proceedings CVPR’89., IEEE Computer Society Conference on IEEE (Veranst.)*, 1989, pp. 618–624 16
- [Samper-González et al. 2018] SAMPER-GONZÁLEZ, Jorge ; BURGOS, Ninon ; BOTTANI, Simona ; FONTANELLA, Sabrina ; LU, Pascal ; MARCOUX, Arnaud ; ROUTIER, Alexandre ; GUILLON, Jérémy ; BACCI, Michael ; WEN, Junhao et al. : Reproducible evaluation of classification methods in Alzheimer’s disease: framework and application to MRI and PET data. In : *Neuroimage* 183 (2018), pp. 504–521 16
- [Schmidt et al. 1996] SCHMIDT, Michael et al. : *Rey auditory verbal learning test: A handbook*. Western Psychological Services Los Angeles, CA, 1996 5
- [Schwarz et al. 2016] SCHWARZ, Christopher G. ; GUNTER, Jeffrey L. ; WISTE, Heather J. ; PRZYBELSKI, Scott A. ; WEIGAND, Stephen D. ; WARD, Chadwick P. ; SENJEM, Matthew L. ; VEMURI, Prashanthi ; MURRAY, Melissa E. ; DICKSON, Dennis W. et al. : A large-scale comparison of cortical thickness and volume methods for measuring Alzheimer’s disease severity. In : *NeuroImage: Clinical* 11 (2016), pp. 802–812 3, 26, 28
- [Shao 1993] SHAO, Jun : Linear model selection by cross-validation. In : *Journal of the American statistical Association* 88 (1993), n. 422, pp. 486–494 23
- [Shen et al. 2012] SHEN, Kai-kai ; FRIPP, Jurgen ; MÉRIAUDEAU, Fabrice ; CHÉTELAT, Gaël ; SALVADO, Olivier ; BOURGEAT, Pierrick ; INITIATIVE, Alzheimer’s Disease N. et al. : Detecting global and local hippocampal shape changes in Alzheimer’s disease using statistical shape models. In : *Neuroimage* 59 (2012), n. 3, pp. 2155–2166 27
- [Shi et al. 2018] SHI, Jun ; ZHENG, Xiao ; LI, Yan ; ZHANG, Qi ; YING, Shihui : Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer’s disease. In : *IEEE journal of biomedical and health informatics* 22 (2018), n. 1, pp. 173–183 59
- [Sonnenburg et al. 2006] SONNENBURG, Sören ; RÄTSCH, Gunnar ; SCHÄFER, Christin ; SCHÖLKOPF, Bernhard : Large scale multiple kernel learning. In : *Journal of Machine Learning Research* 7 (2006), n. Jul, pp. 1531–1565 150

- [Sørensen et al. 2016a] SØRENSEN, Lauge ; IGEL, Christian ; LIV HANSEN, Naja ; OSLER, Merete ; LAURITZEN, Martin ; ROSTRUP, Egill ; NIELSEN, Mads ; INITIATIVE, Alzheimer’s Disease N. ; AUSTRALIAN IMAGING BIOMARKERS the ; AGEING, Lifestyle Flagship S. of : Early detection of Alzheimer’s disease using MRI hippocampal texture. In : *Human brain mapping* 37 (2016), n. 3, pp. 1148–1161 64
- [Sørensen et al. 2016b] SØRENSEN, Lauge ; IGEL, Christian ; PAI, Akshay ; BALAS, Ioana ; ANKER, Cecilie ; LILLHOLM, Martin ; NIELSEN, Mads : Differential diagnosis of mild cognitive impairment and Alzheimer’s disease using structural MRI cortical thickness, hippocampal shape, hippocampal texture, and volumetry. In : *NeuroImage: Clinical* (2016) 64, 66
- [Stone 1974] STONE, Mervyn : Cross-validatory choice and assessment of statistical predictions. In : *Journal of the royal statistical society. Series B (Methodological)* (1974), pp. 111–147 23
- [Sturges 1926] STURGES, Herbert A. : The choice of a class interval. In : *Journal of the american statistical association* 21 (1926), n. 153, pp. 65–66 68, 109
- [Suk et al. 2017] SUK, Heung-Il ; LEE, Seong-Whan ; SHEN, Dinggang : Deep ensemble learning of sparse regression models for brain disease diagnosis. In : *Medical image analysis* 37 (2017), pp. 101–113 xiii, 32, 33, 35, 73, 74, 75, 108, 117, 118, 121, 124, 125, 127, 134, 135, 136, 142
- [Suk et al. 2014] SUK, Heung-Il ; LEE, Seong-Whan ; SHEN, Dinggang ; INITIATIVE, Alzheimer’s Disease N. et al. : Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. In : *NeuroImage* 101 (2014), pp. 569–582 59
- [Tapiola et al. 2008] TAPIOLA, Tero ; PENNANEN, Corina ; TAPIOLA, Mia ; TERVO, Susanna ; KIVIPELTO, Miia ; HÄNNINEN, Tuomo ; PIHLAJAMÄKI, Maija ; LAAKSO, Mikko P. ; HALLIKAINEN, Merja ; HÄMÄLÄINEN, Anne et al. : MRI of hippocampus and entorhinal cortex in mild cognitive impairment: a follow-up study. In : *Neurobiology of aging* 29 (2008), n. 1, pp. 31–38 26
- [Thal et al. 2000] THAL, Dietmar R. ; HOLZER, Max ; RÜB, Udo ; WALDMANN, Guido ; GÜNZEL, Steffen ; ZEDLICK, Dyrk ; SCHÖBER, Ralf : Alzheimer-related τ -pathology in the perforant path target zone and in the hippocampal stratum oriens and radiatum correlates with onset and degree of dementia. In : *Experimental neurology* 163 (2000), n. 1, pp. 98–110 97
- [Thompson et al. 1997] THOMPSON, Paul M. ; MACDONALD, David ; MEGA, Michael S. ; HOLMES, Colin J. ; EVANS, Alan C. ; TOGA, Arthur W. : Detection and mapping of abnormal brain structure with a probabilistic atlas of cortical

- surfaces. In : *Journal of computer assisted tomography* 21 (1997), n. 4, pp. 567–581 [18](#)
- [Tibshirani 1996] TIBSHIRANI, Robert : Regression shrinkage and selection via the lasso. In : *Journal of the Royal Statistical Society. Series B (Methodological)* (1996), pp. 267–288 [21](#)
- [Tolonen et al. 2018] TOLONEN, Antti ; RHODIUS-MEESTER, Hanneke F. ; BRUUN, Marie ; KOIKKALAINEN, Juha ; BARKHOF, Frederik ; LEMSTRA, Afina W. ; KOENE, Teddy ; SCHELTENS, Philip ; TEUNISSEN, Charlotte E. ; TONG, Tong et al. : Data-driven differential diagnosis of dementia using multiclass Disease State Index classifier. In : *Frontiers in Aging Neuroscience* 10 (2018) [143](#)
- [Tong et al. 2017a] TONG, Tong ; GAO, Qinquan ; GUERRERO, Ricardo ; LEDIG, Christian ; CHEN, Liang ; RUECKERT, Daniel ; DISEASE NEUROIMAGING INITIATIVE Alzheimer’s : A Novel Grading Biomarker for the Prediction of Conversion From Mild Cognitive Impairment to Alzheimer’s Disease. In : *IEEE Transactions on Biomedical Engineering* 64 (2017), n. 1, pp. 155–165 [xiii](#), [30](#), [32](#), [35](#), [38](#), [57](#), [64](#), [67](#), [68](#), [73](#), [74](#), [75](#), [78](#), [108](#), [109](#), [117](#), [118](#), [121](#), [124](#), [125](#), [127](#), [134](#), [135](#), [136](#), [142](#)
- [Tong et al. 2017b] TONG, Tong ; GRAY, Katherine ; GAO, Qinquan ; CHEN, Liang ; RUECKERT, Daniel ; INITIATIVE, Alzheimer’s Disease N. et al. : Multimodal classification of Alzheimer’s disease using nonlinear graph fusion. In : *Pattern recognition* 63 (2017), pp. 171–181 [59](#), [143](#)
- [Tong et al. 2017c] TONG, Tong ; LEDIG, Christian ; GUERRERO, Ricardo ; SCHUH, Andreas ; KOIKKALAINEN, Juha ; TOLONEN, Antti ; RHODIUS, Hanneke ; BARKHOF, Frederik ; TIJMS, Betty ; LEMSTRA, Afina W. et al. : Five-class differential diagnostics of neurodegenerative diseases using random undersampling boosting. In : *NeuroImage: Clinical* 15 (2017), pp. 613–624 [143](#)
- [Tong et al. 2013] TONG, Tong ; WOLZ, Robin ; COUPÉ, Pierrick ; HAJNAL, Joseph V. ; RUECKERT, Daniel ; INITIATIVE, Alzheimer’s Disease N. et al. : Segmentation of MR images via discriminative dictionary learning and sparse coding: application to hippocampus labeling. In : *NeuroImage* 76 (2013), pp. 11–23 [57](#), [60](#)
- [Tong et al. 2014] TONG, Tong ; WOLZ, Robin ; GAO, Qinquan ; GUERRERO, Ricardo ; HAJNAL, Joseph V. ; RUECKERT, Daniel ; DISEASE NEUROIMAGING INITIATIVE Alzheimer’s : Multiple instance learning for classification of dementia in brain MRI. In : *Medical image analysis* 18 (2014), n. 5, pp. 808–818 [xiii](#), [30](#), [32](#), [35](#), [39](#), [73](#), [74](#), [75](#), [108](#), [117](#), [118](#), [124](#), [125](#), [134](#), [135](#), [142](#)

- [Trujillo-Estrada et al. 2014] TRUJILLO-ESTRADA, Laura ; DÁVILA, José C. ; SÁNCHEZ-MEJIAS, Elisabeth ; SÁNCHEZ-VARO, Raquel ; GOMEZ-ARBOLEDAS, Angela ; VIZUETE, Marisa ; VITORICA, Javier ; GUTIÉRREZ, Antonia : Early neuronal loss and axonal/presynaptic damage is associated with accelerated amyloid- β accumulation in A β PP/PS1 Alzheimer's disease mice subiculum. In : *Journal of Alzheimer's Disease* 42 (2014), n. 2, pp. 521–541 [89](#), [97](#), [116](#)
- [Tustison et al. 2010] TUSTISON, Nicholas J. ; AVANTS, Brian B. ; COOK, Philip A. ; ZHENG, Yuanjie ; EGAN, Alexander ; YUSHKEVICH, Paul A. ; GEE, James C. : N4ITK: improved N3 bias correction. In : *IEEE transactions on medical imaging* 29 (2010), n. 6, pp. 1310–1320 [17](#), [65](#), [79](#), [90](#), [114](#), [122](#), [130](#)
- [Van Leemput et al. 1999] VAN LEEMPUT, Koen ; MAES, Frederik ; VANDERMEULEN, Dirk ; SUETENS, Paul : Automated model-based tissue classification of MR images of the brain. In : *IEEE transactions on medical imaging* 18 (1999), n. 10, pp. 897–908 [26](#)
- [Vovk et al. 2007] VOVK, Uro ; PERNUS, Franjo ; LIKAR, Botjan : A review of methods for correction of intensity inhomogeneity in MRI. In : *IEEE transactions on medical imaging* 26 (2007), n. 3, pp. 405–421 [17](#), [18](#)
- [Wachinger et al. 2017] WACHINGER, Christian ; BRENNAN, Matthew ; SHARP, Greg C. ; GOLLAND, Polina : Efficient descriptor-based segmentation of parotid glands with nonlocal means. In : *IEEE Transactions on Biomedical Engineering* 64 (2017), n. 7, pp. 1492–1502 [75](#)
- [Wang et al. 2011] WANG, Hongzhi ; DAS, Sandhitsu R. ; SUH, Jung W. ; ALTINAY, Murat ; PLUTA, John ; CRAIGE, Caryne ; AVANTS, Brian ; YUSHKEVICH, Paul A. ; INITIATIVE, Alzheimer's Disease N. et al. : A learning-based wrapper method to correct systematic errors in automatic image segmentation: consistently improved performance in hippocampus, cortex and brain segmentation. In : *NeuroImage* 55 (2011), n. 3, pp. 968–985 [122](#), [130](#)
- [Wang et al. 2015] WANG, Zhiqun ; WANG, Jianli ; ZHANG, Han ; MCHUGH, Robert ; SUN, Xiaoyu ; LI, Kuncheng ; YANG, Qing X. : Interhemispheric functional and structural disconnection in Alzheimer's disease: a combined resting-state fMRI and DTI study. In : *PLoS One* 10 (2015), n. 5, pp. e0126310 [78](#)
- [Wang et al. 2004] WANG, Zhou ; BOVIK, Alan C. ; SHEIKH, Hamid R. ; SIMONCELLI, Eero P. : Image quality assessment: from error visibility to structural similarity. In : *IEEE transactions on image processing* 13 (2004), n. 4, pp. 600–612 [40](#)
- [Wee et al. 2013] WEE, Chong-Yaw ; YAP, Pew-Thian ; SHEN, Dinggang ; INITIATIVE, Alzheimer's Disease N. : Prediction of Alzheimer's disease and mild

- cognitive impairment using cortical morphological patterns. In : *Human brain mapping* 34 (2013), n. 12, pp. 3411–3425 [29](#), [35](#), [108](#), [121](#), [124](#), [125](#), [142](#)
- [Wee et al. 2012] WEE, Chong-Yaw ; YAP, Pew-Thian ; ZHANG, Daoqiang ; DENNY, Kevin ; BROWNDYKE, Jeffrey N. ; POTTER, Guy G. ; WELSH-BOHMER, Kathleen A. ; WANG, Lihong ; SHEN, Dinggang : Identification of MCI individuals using structural and functional connectivity networks. In : *Neuroimage* 59 (2012), n. 3, pp. 2045–2056 [78](#)
- [Weiner et al. 2015] WEINER, Michael W. ; VEITCH, Dallas P. ; AISEN, Paul S. ; BECKETT, Laurel A. ; CAIRNS, Nigel J. ; CEDARBAUM, Jesse ; GREEN, Robert C. ; HARVEY, Danielle ; JACK, Clifford R. ; JAGUST, William et al. : 2014 Update of the Alzheimer’s Disease Neuroimaging Initiative: a review of papers published since its inception. In : *Alzheimer’s & dementia* 11 (2015), n. 6, pp. e1–e120 [22](#)
- [Weiner et al. 2013] WEINER, Michael W. ; VEITCH, Dallas P. ; AISEN, Paul S. ; BECKETT, Laurel A. ; CAIRNS, Nigel J. ; GREEN, Robert C. ; HARVEY, Danielle ; JACK, Clifford R. ; JAGUST, William ; LIU, Enchi et al. : The Alzheimer’s Disease Neuroimaging Initiative: a review of papers published since its inception. In : *Alzheimer’s & Dementia* 9 (2013), n. 5, pp. e111–e194 [22](#)
- [Whitwell et al. 2001] WHITWELL, Jennifer L. ; CRUM, William R. ; WATT, Hilary C. ; FOX, Nick C. : Normalization of cerebral volumes by use of intracranial volume: implications for longitudinal quantitative MR imaging. In : *American Journal of Neuroradiology* 22 (2001), n. 8, pp. 1483–1489 [81](#), [92](#)
- [Wicks et al. 1993] WICKS, David A. ; BARKER, Gareth J. ; TOFTS, Paul S. : Correction of intensity nonuniformity in MR images of any orientation. In : *Magnetic resonance imaging* 11 (1993), n. 2, pp. 183–196 [17](#)
- [Wimo et al. 1997] WIMO, Anders ; LJUNGGREN, Gunnar ; WINBLAD, Bengt : Costs of dementia and dementia care: a review. In : *International Journal of Geriatric Psychiatry* 12 (1997), n. 8, pp. 841–856 [2](#)
- [Winterburn et al. 2013] WINTERBURN, Julie L. ; PRUESSNER, Jens C. ; CHAVEZ, Sofia ; SCHIRA, Mark M. ; LOBAUGH, Nancy J. ; VOINESKOS, Aristotle N. ; CHAKRAVARTY, M M. : A novel in vivo atlas of human hippocampal subfields using high-resolution 3T magnetic resonance imaging. In : *Neuroimage* 74 (2013), pp. 254–265 [90](#)
- [Wolz et al. 2011] WOLZ, Robin ; JULKUNEN, Valtteri ; KOIKKALAINEN, Juha ; NISKANEN, Eini ; ZHANG, Dong P. ; RUECKERT, Daniel ; SOININEN, Hilikka ; LÖTJÖNEN, Jyrki ; DISEASE NEUROIMAGING INITIATIVE Alzheimer’s : Multi-method analysis of MRI images in early diagnostics of Alzheimer’s disease. In : *PloS one* 6 (2011), n. 10, pp. e25446 [16](#), [20](#), [23](#), [26](#), [28](#), [30](#), [35](#), [65](#), [107](#), [142](#)

- [Wörz and Rohr 2006] WÖRZ, Stefan ; ROHR, Karl : Localization of anatomical point landmarks in 3D medical images by fitting 3D parametric intensity models. In : *Medical Image Analysis* 10 (2006), n. 1, pp. 41–58 18
- [Yushkevich et al. 2015] YUSHKEVICH, Paul A. ; AMARAL, Robert S. ; AUGUSTINACK, Jean C. ; BENDER, Andrew R. ; BERNSTEIN, Jeffrey D. ; BOCCARDI, Marina ; BOCCHETTA, Martina ; BURGGREN, Alison C. ; CARR, Valerie A. ; CHAKRAVARTY, M M. : Quantitative comparison of 21 protocols for labeling hippocampal subfields and parahippocampal subregions in in vivo MRI: towards a harmonized segmentation protocol. In : *Neuroimage* 111 (2015), pp. 526–541 89
- [Zhan et al. 2015] ZHAN, Liang ; LIU, Y ; ZHOU, Jiayu ; YE, Jieping ; THOMPSON, Paul M. : Boosting classification accuracy of diffusion MRI derived brain networks for the subtypes of mild cognitive impairment using higher order singular value decomposition. In : *Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on Biomedical Imaging* IEEE (Veranst.), 2015, pp. 131–135 85, 86, 96
- [Zhang et al. 2011] ZHANG, Daoqiang ; WANG, Yaping ; ZHOU, Luping ; YUAN, Hong ; SHEN, Dinggang ; DISEASE NEUROIMAGING INITIATIVE Alzheimer’s : Multimodal classification of Alzheimer’s disease and mild cognitive impairment. In : *Neuroimage* 55 (2011), n. 3, pp. 856–867 59
- [Zhou et al. 2011] ZHOU, Luping ; WANG, Yaping ; LI, Yang ; YAP, Pew-Thian ; SHEN, Dinggang ; (ADNI, Alzheimer’s Disease Neuroimaging I. et al. : Hierarchical anatomical brain networks for MCI prediction: revisiting volumetric measures. In : *PloS one* 6 (2011), n. 7, pp. e21935 108, 129
- [Zhou et al. 2009] ZHOU, Zhi-Hua ; SUN, Yu-Yin ; LI, Yu-Feng : Multi-instance learning by treating instances as non-iid samples. In : *Proceedings of the 26th annual international conference on machine learning* ACM (Veranst.), 2009, pp. 1249–1256 40
- [Zhu et al. 2017] ZHU, Hancan ; CHENG, Hwei ; YANG, Xuesong ; FAN, Yong ; INITIATIVE, Alzheimer’s Disease N. et al. : Metric learning for multi-atlas based segmentation of hippocampus. In : *Neuroinformatics* 15 (2017), n. 1, pp. 41–50 60
- [Zou and Hastie 2005] ZOU, Hui ; HASTIE, Trevor : Regularization and variable selection via the elastic net. In : *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 67 (2005), n. 2, pp. 301–320 21, 110

Acronyms

A

AAL: Automated Anatomical Labeling

ACC: ACCuracy

AD: Alzheimer's Disease

ADAS-Cog: Alzheimer's Disease Assessment Scale-cognitive Subscale

ADNI: Alzheimer's Disease Neuroimaging Initiative

AIBL: Australian Imaging Biomarkers and Lifestyle

ALS: Alternating Least Squares

ANN: Approximate Nearest Neighbors

ANOVA: ANalysis Of VAriances

AUC: Area Under the Curve

C

CA: Cornu Ammonis

CBD: CereBrovascular Disease

CDR-SB: Clinical Dementia Rating–Sum of Boxes

CN: Cognitively Normal

CNN: Convolutional Neural Network

CSF: CerebroSpinal Fluid

CT: Computed Tomography

D

d-MRI: Diffusion Magnetic Resonance Imaging

DG: Dentrate Gyrus

DLB: Dementia with Lewy Bodies

DT: Decision Tree

DTI: Diffusion Tensor Imaging

DWI: Diffusion Weighted Imaging

E

EC: Entorhinal Cortex

EIG: Eigenvalues Decomposition

EN: Elastic Net

EM: Expectation Maximization

eMCI: Early Mild Cognitive Impairment

F

FAQ: Functional Activity Questionnaire

FC: Fully Connected

FD: Frontotemporal Dementia

FDG-PET: FluoroDésoxyGlucose Positron Emission Tomography

fMRI: Function Magnetic Resonance Imaging

G

GBSG: Graph of Brain Structure Grading

GHSB: Graph of Hippocampal Subfields Grading

GSG: Graph of Structure Grading

K

KFCV: K-Fold Cross-Validation

L

LASSO: Least Absolute Shrinkage and Selection Operator

LDS: Low Density Separation

LDA: Linear Discriminant Analysis

lMCI: Late Mild Cognitive Impairment

LOOCV: Leave One Out Cross-Validation

M

MCI: Mild Cognitive Impairment

MD: Mean diffusivity

MGSG: Multiple graph of structure grading

MIL: Multiple Instance Learning

MNI: Montreal Neurological Institute

MMSE: Mini-Mental State Examination

MPBG: Multimodal Patch-Based Grading

MRI: Magnetic Resonance Imaging

MTL: Medial Temporal Lobe

N

NLM: Non Local Mean

MNR: Magnetic Nuclear Resonance

P

PaD: Parkinson's Disease

PBG: Patch-Based Grading

PBL: Patch-Based Labeling

PCA: Principal Component Analysis

PET: Positron Emission Tomography

pMCI: Progressive Mild Cognitive Impairment

p-tau: Phosphorylated-tau

PVE: Partial Volume Error

S

SBG: Sparse-Based Grading

SEN: SENSitivity

SL: Stratum Lacunosum

SLR: Sparse Logistic Regression

SNIFE: Scoring Non local Image Patch Estimator

SM: Stratum Molecular

sMCI: Stable Mild Cognitive Impairment

s-MRI: Structural Magnetic Resonance Imaging

SP: Stratum Pyramidal
SPE: SPEcificity
SPHARM: SPHerical hARMonics
SPM: Statistical Parametric Mapping
SSD: Sum Squared Differences
SSIM: Structural SIMilarity Measure
SR: Stratum Radiatum
STD: STandard Deviation
SVM: Support Vector Machine
SVD: Singular Value Decomposition

R

RAVLT: Rey Auditory Verbal Learning Test
RF: Random Forest
ROI: Region Of Interest

T

T1w: T1 weighted
TBG: Texture-based grading
TBM: Tensor-based morphometry
TIV: Total intracranial volume
t-tau: Total-Tau

O

OASIS: Open Access Series of Imaging Studies
OPAL: Optimized Patchmatch for Accurate Label fusion
OPM: Optimized Patch Match

V

VBM: Voxel-Based Morphometry