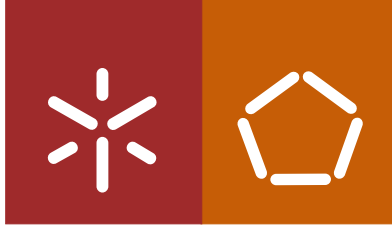


Universidade do Minho
Escola de Engenharia

Patrick Rosa Brandão

**“Computed Aided Diagnostic”
para análise e quantificação de
patologias associadas à demência**



Universidade do Minho
Escola de Engenharia

Patrick Rosa Brandão

**“Computed Aided Diagnostic”
para análise e quantificação de
patologias associadas à demência**

Dissertação de Mestrado
Mestrado Integrado em Engenharia Biomédica
Ramo Eletrónica Médica

Trabalho efetuado sob a orientação do
Professor Doutor Carlos Alberto Baptista da Silva
e supervisão da
Engenheira Maria do Carmo Baptista

outubro de 2014

DECLARAÇÃO

Nome: Patrick Rosa Brandão

Endereço eletrónico: a58550@alunos.uminho.pt

Telemóvel: 912966823

Número do cartão de cidadão: 13929477

Título da Dissertação: “*Computed Aided Diagnostic*” para análise e quantificação de patologias associadas à demência

Orientador: Professor Doutor Carlos Alberto Baptista da Silva

Coorientador: Engenheira Maria do Carmo Baptista

Ano de conclusão: 2014

Designação do Mestrado: Mestrado Integrado em Engenharia Biomédica – Ramo Eletrónica Médica

É AUTORIZADA A REPRODUÇÃO INTEGRAL DESTA DISSERTAÇÃO APENAS PARA EFEITOS DE INVESTIGAÇÃO, MEDIANTE DECLARAÇÃO ESCRITA DO INTERESSADO, QUE A TAL SE COMPROMETE.

Universidade do Minho, ___/___/_____

Assinatura: _____

Agradecimentos

Não posso deixar de expressar um agradecimento especial aos meus orientadores, as pessoas sem as quais este trabalho não seria o mesmo.

Ao Professor Doutor Carlos A. Silva agradeço as suas orientações, conselhos e disponibilização de recursos que foram fundamentais para esta dissertação.

À Dra. Maria do Carmo Baptista agradeço a possibilidade de trabalhar com uma das maiores Clínicas de imagiologia médica do país, a Clínica Dr. Campos Costa, bem como a oportunidade de conhecer os projetos realizados em colaboração com outras universidades.

Por último, um profundo agradecimento a todos os meus amigos, em especial aos meus companheiros de curso, que me ouviram e apoiaram e que permitiram, mesmo que inconscientemente ou por métodos menos convencionais, a resolução de vários dos obstáculos encontrados.

Resumo

O Alzheimer é a forma mais comum de demência, constituindo entre 60% a 80% de todas as demências, com uma incidência aumenta quase exponencialmente com a idade. A neurodegeneração causada afeta vários domínios cognitivos como a memória, o raciocínio, o comportamento e as emoções. Embora as causas e uma cura não sejam ainda conhecidas, um diagnóstico e tratamento precoce ajudam a abrandar a progressão da doença. Métodos neuroimagiológicos providenciam uma forma não invasiva para a análise estrutural e funcional do cérebro, surgindo como uma alternativa de diagnóstico precoce da doença.

O hipocampo é uma das primeiras estruturas a sofrer neurodegeneração, sendo por isso, o foco de um grande número de estudos como forma de distinção entre sujeitos saudáveis e pacientes de Alzheimer. Uma destas investigações foi efetuada por Coupé et al., onde é proposta uma abordagem de detecção de Alzheimer, baseada na distinção de padrões específicos em estruturas anatómicas como o hipocampo e o córtex entorrinal. O método realiza, simultaneamente, a segmentação e a gradação das estruturas de forma a capturar as alterações provocadas pela doença.

Nesta dissertação, é desenvolvido e implementado um método, completamente automático, que permite auxiliar o diagnóstico de Alzheimer e as fases anteriores da doença. Para isso, o algoritmo baseado em *patches* não locais, proposto por Coupé et al., é usado como a base para a criação de uma ferramenta de classificação dos diferentes estados de demência, a partir de imagens RM T1w. É ainda proposta uma alteração ao sistema, sendo a informação das bordas das imagens, obtida pelo filtro Laplaciano, incorporada no cálculo da segmentação e da gradação.

Os métodos implementados foram suficientes para distinguir sujeitos de controlo e pacientes de Alzheimer, obtendo uma exatidão de 85% na classificação da base de dados ADNI-1. No entanto, a previsão de demência, e a determinação da eficácia do novo parâmetro de peso, foram grandemente limitadas pelos problemas originados pelas máscaras do hipocampo usadas.

Índice

Agradecimentos	i
Resumo	IV
Lista de Figuras	VIII
Lista de Tabelas.....	XII
Lista de Abreviaturas, Siglas e Acrónimos.....	XIV
Capítulo 1 Introdução	1
1.1. Motivação.....	1
1.2. Objetivos	2
1.3. Contribuições	2
1.4. Organização da dissertação.....	3
Capítulo 2 Contexto Clínico.....	5
2.1. Doença de Alzheimer	5
2.1.1. Causas e fatores de risco	6
2.1.2. Diagnóstico	7
2.1.3. Tratamento.....	11
2.2. Imagiologia por ressonância magnética.....	15
2.3. O encéfalo	18
2.3.1. Orientação no encéfalo	18
2.3.2. Anatomia	20
2.3.3. Hipocampo.....	23
2.4. Sumário	27
Capítulo 3 Estado da arte.....	29
3.1. Protocolos de segmentação manual do hipocampo e atlas cerebrais.....	30
3.2. Algoritmos de segmentação do hipocampo.....	31
3.2.1. Crescimento regional ou modelos deformáveis.....	32
3.2.2. Métodos baseados aparência e forma	33
3.2.3. Métodos baseados em atlas	34
3.3. Abordagem seguida nesta dissertação.....	38
3.4. Sumário	39

Capítulo 4 Segmentação do hipocampo baseada em patches não locais	41
4.1. Visão geral do método	41
4.2. Pré-processamento	42
4.2.1. Filtragem.....	43
4.2.2. Correção das não-homogeneidades	45
4.2.3. Registo linear para o espaço estereotáxico	47
4.2.4. Normalização das intensidades	48
4.3. Estratégia de procura na base de dados	50
4.3.1. Máscara de inicialização	50
4.3.2. Seleção dos sujeitos.....	51
4.3.3. Pré-seleção de <i>patches</i>	51
4.4. Fusão de <i>labels</i> por média não local	52
4.4.1. Estimador de média não local	52
4.4.2. Adaptação local de h	53
4.5. Sumário	54
Capítulo 5 Segmentação e gradação simultânea de estruturas anatómicas	55
5.1. Da segmentação para a gradação	56
5.2. Base de dados ADNI: 834 imagens de base.....	58
5.3. Pré-processamento	58
5.3.1. Criação de um atlas específico para a população	58
5.3.2. Extração do cérebro.....	61
5.4. SNIFE - <i>Scoring by Nonlocal Patch Estimator</i>	63
5.4.1. Propagação dos <i>labels</i>	63
5.4.2. Segmentação e gradação.....	64
5.4.3. Gradação.....	65
5.5. Esquema de validação	65
5.6. Resultados	66
5.7. Sumário	67
Capítulo 6 Trabalho Experimental	69
6.1. Pré-processamento	71
6.2. Base de dados ADNI1	72

6.3. SNIPE - <i>Scoring by Nonlocal Patch Estimator</i>	75
6.3.1. Linguagem de programação	75
6.3.2. Biblioteca de processamento de imagem.....	75
6.3.3. Propagação dos <i>labels</i>	76
6.3.4. Detalhes de implementação.....	77
6.3.5. Implementação.....	78
6.4. SNIPE baseado no filtro Laplaciano	81
6.4.1. Modificação do cálculo do peso no filtro de média não local	82
6.4.2. Adaptação do peso do SNIPE.....	83
6.4.3. Detalhes de implementação.....	84
6.4.4. Implementação.....	84
6.5. Esquema de validação	86
6.6. Sumário	86
Capítulo 7 Resultados e Discussão.....	89
7.1. Métricas de avaliação.....	89
7.2. Avaliação da segmentação do hipocampo	90
7.3. Estudo volumétrico	95
7.4. Estudo da gradação	100
7.5. Comparação dos biomarcadores do SNIPE	107
7.6. Tempo de computação.....	112
7.7. Sumário	114
Capítulo 8 Conclusão e trabalho futuro	117
8.1. Conclusão	117
8.2. Trabalho Futuro.....	118
Bibliografia	123
Anexo I	130
Anexo II	131

Lista de Figuras

Figura 2.1- Representação de vários núcleos que precessam em torno de um campo magnético, B_0 , com fases aleatórias, não produzindo magnetização transversa. Contudo, ligeiramente mais núcleos estão orientados no mesmo sentido de B_0 , dando origem a uma pequena magnetização longitudinal, M_0 [19].	16
Figura 2.2- Esquemático do plano sagital do encéfalo. Representação do neuraxis, do plano axial e coronal e da orientação [24].	19
Figura 2.3- Principais subdivisões do sistema nervoso central visto num plano sagital. Adaptado de [25].	20
Figura 2.4- Representação dos principais sulcos e fissuras do encéfalo: a) vista lateral; b) localização da ínsula; c) vista em corte do tecido cerebral. Adaptado de [25].	22
Figura 2.5- Aspeto intraventricular do hipocampo. Aqui são visíveis alguns dos tecidos adjacentes ao hipocampo, como o subículo (6) e a fimbria (4), e os seus três segmentos: corpo (1), cabeça (2) e a cauda (3) [26].	24
Figura 2.6- Ilustração da estrutura e das principais ligações da formação hipocampal: hipocampo (três zonas CA3, CA2 e CA1), giro dentado (DG), subículo (Sub), pré-subículo (Pre), para-subículo (Para) e córtex entorrinal (EC) [7].	25
Figura 3.1- Segmentação cerebral baseada em atlas: um atlas é registado com uma nova imagem e os labels são propagados [29].	34
Figura 3.2- Segmentação baseada a múltiplos atlas. As imagens dos atlas são registadas em novas imagem, permitindo a propagação dos seus labels. A segmentação final é obtida por fusão das segmentações individuais [29].	36
Figura 4.1- Pipeline de pré-processamento usado para a construção da biblioteca de imagens. Primeiro, filtragem e correção das não-homogeneidades são feitas no sistema de coordenadas do sujeito. Então, as imagens são registadas para o espaço do atlas MNI-ICBM152. Finalmente, a normalização das intensidades de cada sujeito é feita antes de cortar as imagens à volta das estruturas de interesse [28].	43
Figura 4.2- Correção da não-homogeneidades de imagens de MR T1-w: (a) e (d) vista axial e sagital de imagens não homogéneas; (b) e (e) campo não uniforme estimado pelo método N3; (c) e (f) imagens corrigidas [76].	46

Figura 4.3- Ilustração de uma aplicação de um shear xy a uma imagem RM no plano axial [81].	47
Figura 4.4- Localização dos parâmetros de normalização: a) numa imagem com distribuição bimodal; b) numa imagem com distribuição unimodal [82].	49
Figura 4.5- Visão geral dos diferentes passos envolvidos na segmentação do vóxel xi incluído na máscara de inicialização. O patch Pxi é comparado com todos os patches Pxs, j contidos no volume de procura Vi no interior dos N sujeitos selecionados (neste caso, $N = 3$). O mapa de pesos mostra que os maiores pesos são atribuídos aos patches Pxs, j mais similares a Pxi . Depois da fusão de média não local dos labels ys, j , a estimativa resultante é $v(xi) = 0.994$, pelo que o label final é $L(xi) = 1$ [28].	53
Figura 5.1- Exemplo da execução do SNIPE para um sujeito com MCI. A biblioteca de treino é usada para estimar uma segmentação e um mapa da gradação, que podem ser usados como biomarcadores para o diagnóstico de Alzheimer [60].	57
Figura 5.2- Representação esquemática do algoritmo de construção do modelo: as linhas pontilhadas representam o mapeamento de um vóxel no modelo inicial (modelo 0) para cada sujeito; linhas sólidas representam o mapeamento de sujeitos individuais para o modelo seguinte (modelo 1); linhas tracejadas representam o erro residual de cada vóxel do modelo a cada iteração [86].	61
Figura 5.3- Exemplos de mapas da gradação típicos para cada população de acordo com a idade [60].	64
Figura 6.1- Slice sagital de uma imagem e a respetiva máscara hipocampal no espaço estereotáxico.	72
Figura 6.2- Cronograma ilustrativo das durações das diferentes fases do ADNI.	73
Figura 6.3 - Diagrama representativo da implementação do algoritmo SNIPE.	79
Figura 6.4- Slice sagital de uma imagem sobreposta com união de todas as máscaras da base de dados.	80
Figura 6.5- Diagrama representativo da implementação do algoritmo SNIPE baseado no filtro Laplaciano.	85
Figura 7.1- Slice sagital da pior segmentação obtida pelo SNIPE.	92
Figura 7.2- Slice sagital da pior segmentação obtida pelo SNIPE baseado no operado de Laplace.	93

Figura 7.3- Slice sagital da máscara do hipocampo fornecida pelo ADNI.....	93
Figura 7.4- Estimativa pelo SNIPE implementado do volume do hipocampo em função da idade dos sujeitos. As regressões lineares são apresentadas para uma melhor visualização das tendências globais. O coeficiente de Pearson (R) e os valores-p das regressões lineares são apresentados na legenda.	95
Figura 7.5- Diagrama de caixas das distribuições dos volumes, obtidos pela implementação do SNIPE implementado, para cada população.....	96
Figura 7.6- Estimativa pelo SNIPE-LoG do volume do hipocampo em função da idade dos sujeitos. As regressões lineares são apresentadas para uma melhor visualização das tendências globais. O coeficiente de Pearson (R) e os valores-P das regressões lineares são apresentados na legenda.....	97
Figura 7.7- Diagrama de caixas das distribuições dos volumes, obtidos pelo SNIPE-LoG, para cada população.....	98
Figura 7.8- Diagramas de caixas das distribuições dos volumes, obtidos pelo SNIPE implementado, para as populações sMCI e pMCI obtidas com a informação de conversão até 2011 ou 2014.	99
Figura 7.9- Diagramas de caixas das distribuições dos volumes, obtidos pelo SNIPE-LoG, para as populações sMCI e pMCI obtidas com a informação de conversão até 2011 ou 2014.....	99
Figura 7.10- Estimativa pelo SNIPE implementado da gradação média do hipocampo em função da idade dos sujeitos. As regressões lineares são apresentadas para uma melhor visualização das tendências globais. O coeficiente de Pearson (R) e os valores-P das regressões são apresentados na legenda.	100
Figura 7.11- Diagrama de caixas das distribuições das gradações médias, obtidas pela implementação do SNIPE, para cada população.	101
Figura 7.12- Estimativa pelo SNIPE-LoG da gradação média do hipocampo em função da idade dos sujeitos. As regressões lineares são apresentadas para uma melhor visualização das tendências globais. O coeficiente de Pearson (R) e os valores-P das regressões são apresentados na legenda.....	102
Figura 7.13- Diagrama de caixas das distribuições da gradação média, obtidos pelo SNIPE-LoG, para cada população.....	103

Figura 7.14- Diagramas de caixas das distribuições da gradação média, obtidos pelo SNIPE, para as populações sMCI e pMCI obtidas com a informação de conversão até 2011 ou 2014. 104

Figura 7.15- Diagramas de caixas das distribuições da gradação média, obtidos pelo SNIPE-LoG, para as populações sMCI e pMCI obtidas com a informação de conversão até 2011 ou 2014 104

Figura 7.16- Exemplos de mapa de gradação de acordo com a idade e com a população. . 106

Lista de Tabelas

Tabela 2.1- Descrição das orientações convencionadas para descrever o posicionamento no encéfalo. Os planos indicados correspondem às vistas cujas orientações provocam mudança do slice observado [23], [24].....	19
Tabela 5.1- Resultados obtidos com os diferentes biomarcadores para a diferenciação das diferentes populações. Estes foram obtidos usando a análise discriminante linear com um procedimento LOOCV. Os valores na tabela correspondem à exatidão (Exa),sensibilidade (Sen) e a especificidade (Esp) em percentagem (%) [60].....	66
Tabela 5.2- Resultados obtidos dos melhores biomarcadores para a diferenciação das diferentes populações. Estes foram obtidos usando a análise discriminante linear com um procedimento LOOCV. Os valores na tabela correspondem à exatidão (Exa),sensibilidade (Sen) e a especificidade (Esp) em percentagem (%)	Error! Bookmark not defined.
Tabela 6.1- Detalhes demográficos dos sujeitos usados divididos por população com diagnósticos até 2011. Incluí o número de sujeitos pertencentes (Nº de sujeitos), a percentagem de população masculina (% Masculina), a idade mais o desvio padrão (Idade $\pm \sigma$) e o resultado do MMSE mais o desvio padrão (MMSE $\pm \sigma$)	74
Tabela 6.2- Detalhes demográficos dos sujeitos usados divididos por população com diagnósticos até a atualidade. Incluí o número de sujeitos pertencentes (Nº de sujeitos), a percentagem de população masculina (% Masculina), a idade mais o desvio padrão (Idade $\pm \sigma$) e o resultado do MMSE mais o desvio padrão (MMSE $\pm \sigma$)	74
Tabela 7.1- Valor médio do coeficiente Dice mais o desvio-padrão (Dice $\pm \sigma$) das segmentações do SNIPE implementado. Os resultados são provenientes dos 137 sujeitos CN e 102 sujeitos AD cujas segmentações hipocampais foram fornecidas pelo ADNI.....	91
Tabela 7.2- Valor médio do coeficiente Dice mais o desvio-padrão (Dice $\pm \sigma$) das segmentações do SNIPE-LoG. Os resultados são provenientes dos 137 sujeitos CN e 102 sujeitos AD cujas segmentações hipocampais foram fornecidas pelo ADNI	91
Tabela 7.3- Resultados obtidos para a diferenciação das diferentes populações com os biomarcadores obtidos do hipocampo. São apresentados simultaneamente os resultados obtidos com a implementação do SNIPE e com o SNIPE-LoG. Estes foram obtidos usando a análise discriminante linear com um procedimento LOOCV. Os valores na tabela	

correspondem à exatidão (Exa), sensibilidade (Sen) e a especificidade (Esp) em percentagem (%) e o valor-p (p) do teste McNemar para avaliar a performance dos testes comparados com uma classificação aleatória. É ainda apresentado o Valor-p do teste McNemar para avaliar significância da diferença dos resultados obtidos pelos dois métodos 107

Tabela 7.4- Resultados obtidos para a diferenciação das diferentes populações, criadas com os dados divulgados até 2014, com os biomarcadores obtidos do hipocampo. São apresentados simultaneamente os resultados obtidos com a implementação do SNIPE (SNIPE) e com o SNIPE-LoG (SNIPE-LoG). Estes foram obtidos usando a análise discriminante linear com um procedimento LOOCV. Os valores na tabela correspondem à exatidão (Exa),sensibilidade (Sen) e a especificidade (Esp) em percentagem (%) e o valor-p (p) do teste McNemar para avaliar a performance dos testes comparados com uma classificação aleatória. É ainda apresentado o Valor-p do teste McNemar para avaliar significância da diferença dos resultados obtidos pelos dois métodos 110

Lista de abreviaturas, siglas e acrónimos

AAGP - *American Association for Geriatric Psychiatry*

AChEIS- *Acetilcolinesterase*

A β - β -amilóide

AD- *Alzheimer*

ADAS-Cog - *Alzheimer's Disease Assessment Scale – Cognitive*

ADNI - *Alzheimer's Disease Neuroimaging Initiative*

ANIMAL- *Automatic Nonlinear Image Matching and Anatomical Labeling*

BEaST- *Brain extraction based on nonlocal segmentation technique*

CA- *Corno de Amon*

CAD- *Computed Aided Diagnostic*

CN-*Controlo*

DARTEL- *Diffeomorphic Anatomical Registration Through Exponential Lie Algebra*

EC- *Cortex entorrinal*

EM- *Expectation-maximization*

ChEI- *Colinesterase*

FN- *Falsos negativos*

FP- *Falsos positivos*

FMASH- *Fast marching for automated segmentation of the hippocampus*

HC- *Hipocampo*

ITK- *Insight Toolkit*

LCR - *Líquido Cefalorraquidiano*

LDA- *Linear Discriminant Analysis*

LDDMM- *Large Deformation Diffeomorphic Metric Mapping*

LoG- *Filtro Laplaciano*

LOOCV- *Leave-one-out cross-validation*

MCI - *Mild Cognitive impairment*

MINC- *Medical Image NETCDF*

MIS - *Memory Impairment Screen*

MMSE - *Mini-Mental Status Examination*

MNI- *Montreal Neurological Institute*

MoCA - *Montreal Cognitive Assessment*

MRF- *Markov random field*

MRI- *Magnetic Resonance Imaging*

NIA - *National Institute on Aging*

NINCDS-ADRDA - *National Institute of Neurologic, Communicative Disorders and Stroke/Alzheimer's disease and Related Disorders Association*

N3- *Nonparametric Nonuniform intensity Normalization*

PET- *Positron emission tomography*

pMCI- *Progressive Mild Cognitive impairment*

QDA- *Quadratic Discriminant Analysis*

RLS- *Recursive Least Square*

RM- *Ressonância Magnética*

SACHA- *Segmentation Automatisée Compétitive de l'Hippocampe et de l'Amygdale*

SDQ- *Soma dos quadrados das diferenças*

sMCI- *static Mild Cognitive impairment*

SNIFE - *Scoring by Nonlocal Patch Estimator*

SNIFE-LoG - *SNIFE baseado no filtro Laplaciano*

STAPLE- *Simultaneous truth and Performance Level Estimation*

TAC- *tomografia axial computadorizada*

VN- *Verdadeiros negativos*

VP- *Verdadeiros positivos*

Capítulo 1 Introdução

Este capítulo começa por descrever a motivação e os objetivos do trabalho. São ainda apresentadas as contribuições resultantes do projeto, terminando com a descrição da organização da dissertação.

1.1. Motivação

É estimado que, em 2010, existiram 35.6 milhões de pessoas a viver com demência em todo o mundo e é esperado que este número suba para 115.4 milhões em 2050, se os esforços para prevenir a doença falharem. O maior aumento é esperado nos países menos desenvolvidos, onde a taxa de aumento anual em adultos com mais de 60 anos é maior que nos países desenvolvidos [1].

A doença de Alzheimer é, seguramente, a forma mais comum de demência constituindo entre 60% a 80% de todas as demências e a sua incidência aumenta quase exponencialmente com a idade [2]. Atualmente, nenhum fator único é atribuído à causa da doença, pensando-se que esta se deve a uma combinação de fatores genéticos e ambientais que afetam o cérebro ao longo de um grande período de tempo [2], [3]. Da mesma forma, nenhuma cura é ainda conhecida, mas um diagnóstico e tratamento precoce ajudam a abrandar a progressão da doença [2]–[4].

Um diagnóstico definitivo de Alzheimer exige uma análise microscópica do tecido cerebral do paciente o que, por ser um teste demasiado invasivo, impossibilita-o de se tornar um exame viável. Na maioria dos casos, o Alzheimer é diagnosticado clinicamente. Inicialmente é feito um levantamento do histórico do paciente e do que o levou à consulta, seguindo-se de testes físicos e mentais. Este tipo de abordagem sofre de grande subjetividade e pressupõem um certo avanço da doença [2].

Estudos neuroimagingológicos providenciaram informações vitais e provaram que o processo degenerativo inicia-se anos antes da existência de qualquer sintoma [1]. Estes surgem como uma alternativa para um diagnóstico precoce e não invasivo da doença [2], [3]. A ressonância magnética oferece um método não invasivo para a análise estrutural e

funcional do cérebro sem recorrer a radiações ionizantes. Isto permite a realização de estudos longitudinais sem prejudicar a saúde dos pacientes [5].

1.2. Objetivos

O principal objetivo da dissertação é o desenvolvimento e implementação de um método completamente automático que permita auxiliar o diagnóstico, preferencialmente precoce, de um paciente com Alzheimer. Para isso, será implementado um algoritmo de segmentação e gradação automática do hipocampo usando o algoritmo de *patches* não locais propostos por Coupé. A posterior validação será efetuada com as imagens fornecidas pelo ADNI, a mesma base de dados utilizada por Coupé, de forma a permitir comparações diretas dos métodos.

Será ainda proposta uma modificação do algoritmo de forma a incluir informação das arestas das imagens no processamento efetuado. Como as fronteiras do hipocampo são regiões de grande variabilidade, abordagens de *patches* não locais acabam por atribuir valores menos distintivos aos vóxeis das bordas. O objetivo desta alteração é melhorar a medida de comparação destas zonas do hipocampo.

1.3. Contribuições

Nesta dissertação foi desenvolvido um sistema de fusão de *labels*, com uma nova forma de pesagem, para a segmentação e gradação do hipocampo em imagens de RM que, no conhecimento do autor, pode corresponder a contribuições originais.

Inspirado pelo filtro de média não local proposto por Bhujle et al. [6], o cálculo das similaridades do SNIPE foi alterado de forma a ter em conta a informação das arestas dos *patches* comparados, sendo esta obtida por aplicação do filtro Laplaciano. A nova forma de pesagem foi sugerida com o intuito de dar maior importância a *patches* com grande similaridade de bordas, melhorado assim, a classificação dos vóxeis das extremidades do hipocampo.

1.4. Organização da dissertação

A dissertação foi dividida em oito capítulos, ao longo dos quais se apresentam os fundamentos teóricos necessários, o trabalho experimental implementado, os resultados obtidos e as conclusões retiradas.

O Capítulo 1 apresenta apenas as motivações e objetivos da realização deste trabalho. No Capítulo 2 são introduzidos alguns dos mecanismos conhecidos da neurodegeneração causada pelo Alzheimer bem como alguns dos fatores de risco, métodos de diagnóstico e tratamentos mais comuns. São ainda abordados alguns conceitos teóricos da imagiologia por ressonância magnética, bem como dadas algumas noções anatómicas do encéfalo, com especial destaque para o hipocampo.

Alguns dos principais métodos de diagnóstico de demência a partir de MRI são mencionados no Capítulo 3. Devido à eficácia dos resultados reportados pelas análises do hipocampo, algumas das principais abordagens de segmentação hipocampal são também apresentados.

O Capítulo 4 e o Capítulo 5 descrevem, respetivamente, o método de segmentação e de gradação proposto por Coupé et al. que serão usados como base do sistema de *Computer Aided Diagnostic* (CAD) implementado. Para além dos algoritmos de *patches* não locais, são também descritos os métodos de pré-processamento necessários bem como os esquemas de validação usados.

O trabalho experimental realizado é apresentado no Capítulo 6. Aqui é descrita a base de dados e a biblioteca de processamento de imagem usada, bem como as principais opções tomadas na implementação do algoritmo de Coupé. É ainda descrita uma nova forma de calcular a similaridade entre *patches*, onde é incluída a informação das bordas.

No Capítulo 7 reportam-se os resultados obtidos em termos de qualidade de segmentação do hipocampo e de capacidade de diagnóstico dos biomarcadores calculados.

Por fim, o Capítulo 8 refere as principais conclusões da dissertação e o possível trabalho futuro.

Capítulo 2 Contexto Clínico

Neste capítulo são descritos os aspectos mais gerais conhecidos sobre a doença de Alzheimer. Detalhes sobre causas e fatores, diferentes formas de diagnóstico e tratamento são também discutidos. Devido ao crescimento da importância da imagiologia por ressonância magnética no acompanhamento da doença, a técnica de RM, e os seus conceitos básicos, são também apresentados. Por fim, uma descrição anatômica do encéfalo é feita, mencionando as orientações convencionadas e algumas das estruturas mais importantes. Especial importância foi dada ao hipocampo devido à sua forte correlação com o Alzheimer [7].

2.1. Doença de Alzheimer

Demência é o nome mais comum para síndromes clínicas responsáveis por uma deterioração progressiva do cérebro [8]. Esta afeta vários domínios cognitivos como a memória, o raciocínio, o comportamento e as emoções, de tal forma que compromete a atividade social e profissional [8], [9]. A doença de Alzheimer, ou simplesmente Alzheimer, é o tipo mais comum de demência levando, por isso, ao uso da palavra “demência” indistintamente com Alzheimer. No entanto, existem ainda outros tipos de demência, como a demência vascular e a frontotemporal [2].

A doença de Alzheimer foi chamada a partir do Dr. Alois Alzheimer. Em 1906, o Dr. Alzheimer notou algumas mudanças no tecido do cérebro de uma mulher que tinha morrido com uma doença mental. Anteriormente à sua morte, os sintomas apresentados incluíam perda de memória, problemas na fala e comportamento imprevisível. Ao examinar o cérebro, o Dr. Alzheimer encontrou vários aglomerados anormais, agora chamados de placas amilóides, e emaranhados de fibras, atualmente conhecidos como emaranhados neurofibrilares [10]. As placas amilóides são depósitos densos e insolúveis de peptídios β -amilóide ($A\beta$) acumulados fora das células, enquanto os emaranhados neurofibrilares são agregados insolúveis e intracelulares de proteínas tau. Ainda hoje, estas são as duas principais características fisiológicas da doença [9].

Embora ainda não seja conhecido o que desencadeia o processo da doença, é estimado que os danos ao cérebro começam entre 10 a 20 anos antes que qualquer problema seja evidente. À medida que cada vez mais placas e emaranhados se formam numa área específica do cérebro, neurónios saudáveis começam a trabalhar menos eficientemente. Estes começam por perder funcionalidade e capacidade de comunicar com neurónios vizinhos, acabando eventualmente por morrer. O processo inevitavelmente alastra-se para as zonas cerebrais mais próximas. À medida que o número de neurónios mortos aumenta, regiões do cérebro afetadas começam a atrofiar. Em fases finais da doença, grande parte do cérebro diminui significativamente de volume [10].

2.1.1. Causas e fatores de risco

Todas as formas de demência resultam da morte de células nervosas e/ou perda de capacidade de comunicação entre essas células. O cérebro humano é um órgão muito complexo e vários fatores podem interferir no seu funcionamento. Muitos anos de investigação desvendaram vários destes fatores, mas o desenvolvimento completo das demências não é ainda conhecido [10].

Acredita-se que entre 1% a 5% dos casos de Alzheimer devem-se a mutações genéticas, enquanto a maior parte é atribuída a fatores ambientais e as suas interações com suscetibilidade genética. A idade é o maior determinante de demência, sugerindo que processos biológicos relacionados com o envelhecimento podem estar envolvidos na patogénese da doença. Outros fatores de risco são baseados em várias hipóteses: suscetibilidade genética, vascularização, nutrição e outros (como por exemplo, fatores tóxicos ou inflamatórios). Enquanto o papel dos fatores genéticos e vasculares é fortemente suportado por investigações epidemiológicas e imagiológicas, as evidências para os restantes fatores são controversas ou insuficientes [9].

O histórico familiar foi um dos primeiros fatores de risco identificado. Indivíduos que têm pelo menos um familiar de primeiro grau com demência têm um maior risco de desenvolvimento da doença. Até o momento, apenas o alelo ApoE ϵ 4 foi estabelecido como um gene suscetível para Alzheimer, no entanto, com a era de estudos de associação de

genoma, novos genes candidatos têm sido identificados, mas com uma associação menos consistente [1], [10].

Vários estudos têm demonstrado que elevada pressão sanguínea a meado da vida pode aumentar o risco de Alzheimer em idades mais avançadas. No entanto, um declínio da pressão sanguínea é observado alguns anos precedentes ao diagnóstico. Uma associação similar foi detetada com os valores de colesterol. Tem sido especulado que reduções da pressão sanguínea e dos níveis de colesterol podem refletir um processo neurodegenerativo em andamento. Contudo as relações entre estes fatores são complexas e podem refletir fenômenos relacionados com a fisiologia do envelhecimento ou mudanças no estilo de vida [1].

Existe um aumento de evidências que sugerem que a diabetes tipo 2 está associado com o aumento do risco do desenvolvimento de Alzheimer ou outras demências vasculares. Embora nem todos os estudos confirmem esta relação, existem vários mecanismos pelos quais a diabetes pode influenciar o risco de demência, como a hiperinsulinemia, doença cerebrovascular e o aumento do *stress* oxidativo [1].

Vários estudos recentes descobriram um aumento significativo no risco de declínio mental associado ao tabaco. Fumadores estão mais suscetíveis a doenças vasculares, o que pode levar ao aumento do risco de demência. Estudos também descobriram a ingestão de grandes quantidades de álcool aumentam o risco de Alzheimer embora, aparentemente, pessoas que consomem quantidades moderadas de bebidas alcoólicas têm menor risco de desenvolvimento de demência do que pessoas que se absterem completamente. Uma interação entre suscetibilidade e o estilo de vida também foi observado: pessoas com ApoE ϵ 4 e um estilo de vida menos saudável podem ser especialmente vulneráveis ao desenvolvimento de Alzheimer [1], [10].

2.1.2. Diagnóstico

Em 1984, os critérios de diagnóstico criados pelo *National Institute of Neurologic, Communicative Disorders and Stroke/Alzheimer's disease and Related Disorders Association* (NINCDS-ADRDA), eram baseados na opinião clínica do médico sobre a causa dos sintomas do paciente, relatórios do paciente e dos seus familiares, o resultado de testes cognitivos e

uma avaliação neurológica geral. Estes critérios têm sido bem-sucedidos e têm sido largamente usados em vários ensaios e investigações clínicas, mostrando uma sensibilidade de 81% e uma especificidade de 70% [9], [11]. Contudo, o aumento do conhecimento das manifestações clínicas e biológicas do Alzheimer determinaram a necessidade de uma revisão dos critérios. Em 2011, o *National Institute on Aging* (NIA), e a *Alzheimer's Association* recomendaram novos critérios de diagnóstico. As principais mudanças assentam no conceito que o processo patofisiológico do Alzheimer inicia-se anos antes ao diagnóstico de demência e incorporam biomarcadores que indicam a presença ou ausência da doença [9].

O diagnóstico em cenário clínico é normalmente dividido em duas fases: *screening*, onde é formulada uma hipótese de diagnóstico, seguido por uma confirmação de diagnóstico. Durante a fase de *screening*, o objetivo principal, tal como anteriormente, é a recolha detalhada de informação, não só a partir do paciente, mas também a partir seus familiares e das pessoas que cuidam dele. Esta inclui o histórico médico (presença de doenças que podem causar encefalopatias, doenças psiquiátricas, danos cerebrais ou outras doenças neurológicas), medicação, historial familiar de demência e alterações de atividades básicas do dia-a-dia. Uma avaliação sistemática do funcionamento geral cognitivo é necessário, embora não exista uma forma rápida de o fazer. As diretrizes da *American Academy of Neurology* sugerem o uso do *Mini-Mental Status Examination* (MMSE), e o *Memory Impairment Screen* (MIS). Recentemente, o *Montreal Cognitive Assessment* (MoCA), foi desenvolvido como ferramenta para filtrar pacientes onde existe suspeita de um ligeiro declínio cognitivo mas que geralmente têm resultados normais no MMSE [9].

Depois do *screening*, a segunda fase consiste num exame neurológico e uma avaliação neurofisiológica e comportamental. A presença de síndrome de Parkinson pode sugerir demência de Lewy Body, enquanto reflexos assimétricos dos tendões, ou outros sinais de laterização, podem insinuar um componente vascular. Embora não existam evidências que testes sanguíneos sejam suficientes para a avaliação de demência, estes podem ajudar a excluir outras patologias. Neuroimagiologias estruturais, como a tomografia computadorizada ou a ressonância magnética são recomendadas, embora não essenciais, como forma de exclusão de tumores ou hematomas subdurais [9]. Este conjunto de critérios

tem a vantagem de poder ser usado por instituições de saúde sem acesso a técnicas de imagiologia avançadas ou análise do líquido cefalorraquidiano - LCR [5].

Ainda que a exatidão do diagnóstico tenha aumentado, um diagnóstico sistemático continua a ser problemático, sendo este largamente baseado na exclusão de causas secundárias para a perda de memória. Num estado prematuro de déficit cognitivo, frequentemente classificado como *mild cognitive impairment* (MCI), discernir entre processo degenerativo da doença de “envelhecimento normal” pode ser difícil. Para além disso, a progressão patológica é lenta, e existe variabilidade nos desempenhos das medidas clínicas, sendo difícil monitorizar mudanças de uma forma eficiente. Uma vez que as terapias atuais são mais eficientes quando mais cedo forem aplicadas, existe a necessidade de um biomarcador capaz de diagnosticar precocemente a doença [5].

Critérios para um bom biomarcador foram propostos pelo grupo *The Ronald and Nancy Reagan Research Institute of the Alzheimer’s Association and the National Institute on Aging Working Group*, em 1998. Segundo estas diretrizes, um biomarcador para Alzheimer deve detetar manifestações da neuropatologia fundamental e ser validados em casos confirmados. A sua sensibilidade deve exceder os 80% e a sua especificidade na distinção de outras doenças deve ser superior a 80%. Idealmente, um biomarcador deve ser fiável, reproduzível, não-invasivo, de execução simples e económico. Por fim, no caso do Alzheimer, este deve ser capaz de diagnosticar a doença o mais cedo possível [5].

Uma das fontes de biomarcadores de Alzheimer mais promissora é o líquido cefalorraquidiano. Um único extrato de tecido epitelial separa o cérebro do LCR permitindo um troca virtualmente desimpedida de moléculas entre os dois componentes. Assim, os seus biomarcadores refletem os processos patogénicos centrais do cérebro. Além disso, o LCR é acessível através de punções lombares. Nos candidatos para biomarcadores do LCR estão incluídos a tau total (T-tau), como marcador para a degeneração neuronal e a concentração de A β , como marcador da formação de placas [5].

Vários estudos encontraram um declínio no A β e um aumento do tau total em pacientes com Alzheimer. A redução de A β pode chegar aos 50% comparado com controlos da mesma idade, pensando-se ser o resultado da deposição de A β sob forma de placas, deixando menos A β para ser difundido para o LCR. A T-tau no LCR reflete a intensidade dos

danos neuronais e axoniais e existe em quantidades 2 a 3 vezes superiores em pacientes com demência. No entanto, a quantidade de proteína tau pode ser aumentada devido a um qualquer ferimento cerebral grave, como no caso de acidente vascular cerebral ou trauma. Usando uma combinação das concentrações de A β e T-tau, maiores sensibilidades e especificidades podem ser alcançadas. Apesar do grande potencial na área da investigação, os biomarcadores de LCR não são recomendados no ambiente de diagnóstico clínico, uma vez que a padronização destes marcadores é limitada a hospitais universitários e obrigam os pacientes a passar, potencialmente em várias ocasiões, por um procedimento invasivo e, ocasionalmente, doloroso [5], [9].

A imagiologia com a PET mostra o metabolismo do cérebro, e reflete padrões de neurodegradação. Uma redução metabólica pode ser descrita como um critério diagnóstico de Alzheimer, existindo alguns estudos que revelaram hipometabolismo cerebral vários anos antes do aparecimento de qualquer sintoma clínico. Recentemente foram desenvolvidos marcadores capazes de ligar-se a fibras amiloides do cérebro, em quantidades suficientes para empregar a PET. No entanto, a utilização da PET obriga ao uso de materiais radioativos, com tempos de degradação muito curtos, que necessitam de ser criados no local, tornando PET dispendioso e limitado a centros especializados [5], [9].

Os resultados das sensibilidades e especificidades de cada um dos marcadores mencionados estão apresentados na Tabela 2

Tabela 2 - Valores de especificidade e sensibilidade para biomarcadores de Alzheimer [5], [9]

Biomarcadores de deposição de Aβ	Especificidade	Sensibilidade
A β no LCR	90%	86%
PET com marcadores de amiloides	56.2%	93.5%
Biomarcadores de danos neuronais	Especificidade	Sensibilidade
Tau no LCR	90%	81%
PET	74%	78.7%

Em contraste com outras técnicas, a ressonância magnética oferece um método não invasivo para a análise estrutural e funcional do cérebro sem recorrer a radiações ionizantes. Isto permite a realização de estudos longitudinais sem prejudicar a saúde dos pacientes.

Como estudos de ressonância magnética requerem uma imobilização completa por vários minutos, e alguns scans funcionais requerem que o paciente se foque na realização de tarefas complexas, a realização deste tipo de estudos imagiológicos, em casos avançados de Alzheimer, é muito difícil. Habitualmente, a maioria das análises foca-se em casos moderados de déficit cognitivo. Em alguns casos, estes exames são feitos em fases irreversíveis da doença, no entanto, estes têm vindo a provar que alterações patológicas ocorrem muito antes de qualquer declínio cognitivo aparente. Assim, estudos que analisem as diferenças entre pessoas que eventualmente desenvolvem Alzheimer e os restantes participantes fornecem informação sobre as causas e a fisiologia da doença [5].

2.1.3. Tratamento

As técnicas terapêuticas de Alzheimer são normalmente divididas em três categorias: estratégia de prevenção, tratamento dos sintomas e terapias modificadoras da doença [9].

Atualmente existe uma longa lista de fatores que podem reduzir ou atrasar a progressão da doença, mas a eficácia da prevenção está ainda por ser provada irrefutavelmente. Não existe nenhuma cura conhecida, mas novos tipos de tratamentos modificadores da doença estão sob investigação. Para além disso, intervenções não farmacêuticas têm sido recentemente utilizadas como forma de controlo sintomático [9].

A neuropatologia do Alzheimer é caracterizada por uma perda prematura de neurónios colinérgicos, levando a uma diminuição da transmissão colinérgica. Esta está envolvida em vários aspetos da memória e da atenção. Inibidores da enzima acetilcolinesterase (AChE), aumentam o nível de acetilcolina no cérebro, o que leva a um melhoramento da memória. Desde a introdução da primeira colinesterase (ChEI) em 1997, estes agentes têm sido considerados como primeira linha de terapia farmacêutica para estágios leves ou moderados de Alzheimer [9].

Os resultados dos ensaios clínicos controlados com placebo das ChEIs demonstrou que o tratamento obtém melhoramentos entre 1.4 e 3.9 pontos na escala Alzheimer's Disease Assessment Scale – Cognitive (ADAS-Cog), em 6 meses. Uma mudança de 4 pontos é considerada clinicamente significativa para pacientes com demência leve e moderada. Em adição aos seus efeitos na perceção, estes agentes também demonstraram efeitos benéficos

no comportamento, atividades diárias e nas funções globais do paciente [9]. Quatro ChEIs estão atualmente disponíveis: tacrina, donepezila, rivastigmina e galantamina. A tacrina, o primeiro fármaco aprovado, não é normalmente usada devido à sua pobre tolerabilidade e biodisponibilidade oral (baixa capacidade da dose administrada atingir a circulação sistêmica) [9], [12]. A donepezila foi aprovada a meio da década de 90 obtendo respostas até 52 semanas. Depois da sua utilização a performance do paciente retorna à anterior do tratamento e os seus efeitos secundários incluem náuseas, vômitos, gastrite e diarreia. A rivastigmina é um inibidor enzimático pseudo-irreversível que se dissocia da enzima lentamente. Existem dois tipos de administração possível: oral e adesivo transdérmico. Os efeitos secundários são aproximadamente iguais aos da donepezila, apesar dos sintomas gastrointestinais serem pelo menos 3 vezes menos prominentes com o adesivo. A terapia com galantamina começa com 4 mg duas vezes ao dia, com incrementos de 4 mg, até um máximo de 12 mg, se tolerado. Alternativamente, também está disponível como uma fórmula de libertação controlada que pode ser tomada uma vez por dia. Todos os AChEIs podem influenciar o ritmo cardíaco, se o paciente já tiver um distúrbio cardíaco, pelo que um eletrocardiograma antes do tratamento é recomendado. As ChEIs podem ainda ter um efeito prejudicial em problemas respiratórios, como doença pulmonar obstrutiva crônica e asma, ou em doenças gastrointestinais, como úlceras gástricas. Estes agentes são apenas recomendados em fases leves ou moderadas da doença [9].

Uma das hipóteses das causas de Alzheimer assenta numa disfunção na neurotransmissão por glutamato, que pode provocar danos e a morte em células nervosas devido a uma excitação celular excessiva. A memantina ataca os recetores de glutamato, reduzindo o excesso de excitação. Ensaio clínicos com memantina reportam ligeiras eficácias na manutenção do nível funcional em pacientes com demência severa, sendo licenciado para o tratamento de pessoas com demência moderada e severa. A dose inicial é de 5 mg, e pode ser aumentada 5 mg a cada semana, até os 20 mg. Efeitos secundários são muito incomuns, mas podem incluir inquietação, hiperexcitação e cansaço. Existem provas clínicas dos benefícios do uso de memantina em conjunto com AChEIs em estados avançados de Alzheimer [9].

Uma vez que o beta-amilóide (A β), é considerado como tendo um papel supremo no desenvolvimento de Alzheimer, varias estratégias de investigação têm sido levadas no sentido da alteração da bioquímica do A β no cérebro, através da interferência na sua formação ou na sua deposição. A proteína precursora amilóide pode ser processada por dois processos distintos: de forma não amiloidogénica pela α -secretase e de forma amiloidogénica pela β -secretase seguida pela γ -secretase. Assim, algumas terapias usam inibidores, com o objetivo da redução da produção de β e γ secretases. Apesar da promissora plausibilidade biológica destes compostos, os resultados dos ensaios clínicos têm sido dececionantes [9].

Novas estratégias imunológicas procuram a remoção de A β do cérebro. Abordagens iniciais, baseadas na imunização com fragmentos A β , obtiveram resultados excelentes em ratos transgénicos, mas mostraram-se menos promissores em humanos. Uma análise pós-morte dos participantes do ensaio clínico confirmou que a terapia resultou numa redução significativa de A β no cérebro, no entanto, não houve qualquer provas de qualquer benefício clínico. O ensaio foi interrompido devido ao facto dos pacientes desenvolverem meningoencefalite assética, induzida por uma ativação citotóxica de células T. Atualmente, imunoterapias têm usado peptídios mais curtos para simular secções imunorreativas da A β , para evitar reações inflamatórias [9], [13].

Imunoterapia para Alzheimer tem encontrado alguma crítica. Existe uma dificuldade em desenvolver abordagens capazes de introduzir concentrações suficientes de anticorpos no cérebro. Os resultados dos ensaios a decorrer são ansiosamente aguardados [14]–[16]. Não obstante o facto de uma vacina conseguir remover A β do cérebro, o debate sobre o seu benefício clínico continua. Emaranhados neurofibrilares são outra característica importante da patologia, embora tratamentos direcionados para a proteína tau recebam muito menor atenção do que as terapias amiloides [9].

Terapias não-farmacêuticas para o atraso da progressão da doença têm sido de grande interesse público ao longo dos últimos 10 anos. Existem várias abordagens para intervenções psicossocial e de treino com diferentes provas de eficácia em pacientes com Alzheimer. Estas podem ser classificadas de acordo com o objetivo do tratamento e podem ser orientadas ao comportamento, às emoções, à estimulação e ao treino cognitivo. Estas obtiverem

resultados modestos, quando usadas isoladamente, pelo que a *American Association for Geriatric Psychiatry* (AAGP) propôs um tratamento que combina terapias farmacêuticas e não-farmacêuticas. Até a data, a literatura sobre intervenções psicossociais é muito vasta pelo que, nesta dissertação, apenas se mencionam as intervenções mais importantes [9].

A terapia orientada ao comportamento é usada para modificar costumes disfuncionais, usando técnicas que aumentam ou diminuem a frequência de comportamentos através do uso de reforço e punição. Esta terapia é particularmente útil para reduzir problemas típicos, como incontinência e divagação. Intervenções orientadas à estimulação incluem atividades recreativas, terapias com arte, música ou animais e outras atividades que maximizem o bem-estar dos pacientes. A estimulação melhora, modestamente, comportamentos e o humor, mas o maior efeito é na mudança da rotina diária. Abordagens direcionadas à emoção incluem psicoterapias de apoio, terapia de reminiscência, terapia de validação e treinos cognitivos. A Psicoterapia de apoio não recebeu ainda estudos científicos formais, mas pode ser usada para falar do problema de perda, em fases iniciais de Alzheimer e ajudar pacientes a ajustar-se ao novo estilo de vida imposto pela doença. Terapia de reminiscência permanece controversa, individualmente ou em grupo, esta tenta recordar o paciente de eventos passados com ajuda externa, como fotografias, objetos domésticos, música ou outros itens familiares. Terapia de validação baseia-se numa relação de empatia entre o paciente e o terapeuta. Este ouve o paciente e examina a sua percepção da realidade em ordem a criar contatos emocionais relevantes. O objetivo é estimular o paciente em assumir um papel social, promover a comunicação verbal e encorajar interação social. O treino cognitivo envolve um treino guiado num conjunto de tarefas padrão, projetadas para estimular funções cognitivas específicas (memória, atenção e capacidade de resolução de problemas). O principal pressuposto é que a prática pode melhorar ou manter a funcionalidade num certo domínio, generalizando os efeitos do treino para a vida diária. Tem sido demonstrado que o treino cognitivo melhora a funcionalidade cognitiva, mas os efeitos são transientes e frequentemente acompanhados por efeitos negativos ligados à frustração [5], [9].

2.2. Imagiologia por ressonância magnética

A imagiologia por ressonância magnética (MRI) fornece uma forma não invasiva de investigar anatomia e fisiologia interna. Esta explora o fenómeno de ressonância magnética (RM) nuclear, onde núcleos atômicos expostos a fortes campos magnéticos absorvem e reemitem ondas eletromagnéticas, a uma frequência específica. A inexistência de efeitos adversos conhecidos torna-a um dos métodos de eleição para diagnóstico e estudos científicos em animais e humanos [17].

A ressonância magnética nuclear assenta na propriedade de núcleos com prótons ou neutrões desemparelhados, conhecida como momento angular de *spin*. O *spin* de um núcleo pode ser considerado como o resultado do movimento rotacional do núcleo em torno do seu próprio eixo o que, por se tratar de cargas elétricas em movimento, origina um campo magnético. A ressonância magnética utiliza o comportamento dos núcleos de hidrogénio, pois este é o elemento mais abundante do corpo humano [18].

Na presença de um campo magnético o núcleo comporta-se como um pequeno íman, havendo uma interação de campos magnéticos. Quando se aplica um campo externo, B_0 , a existência de um momento angular faz com que o elemento precesse (mude o eixo de rotação) em redor do eixo desse campo, a uma frequência dependente do elemento [18]. A frequência à qual o núcleo precessa em torno do campo magnético é chamada de frequência de Larmor, e pode ser calculada pela Equação 2.1.

$$W_L = \gamma B_0 \quad (2.1)$$

onde W_L é a frequência de precessão, B_0 a intensidade do campo magnético e γ é o rácio giromagnético do elemento [19].

À medida que o núcleo precessa em torno de B_0 , o seu momento magnético altera-se à frequência de Larmor, produzindo um campo magnético oscilatório. A matriz de oscilações magnéticas gerada por todos os núcleos da amostra em estudo pode ser detetada com uma bobina detetora de radiofrequências. No entanto, o sinal total seria nulo a não ser que um número macroscópico de núcleos da amostra precesse em sincronia. Isto pode ser conseguido por excitação com rádio frequência [19].

O campo magnético de um núcleo é uma função complexa de espaço, mas pode ser especificada por um momento magnético, μ . O campo magnético de todos os núcleos de um dado volume de tecido pode, da mesma forma, ser especificado pela soma de vetores dos seus momentos magnéticos. A soma é chamada de magnetização nuclear, e é denotada M . O componente M é perpendicular ao campo B_0 , rodando à frequência de Larmor, ω_L , à medida que os núcleos precessam. Isto produz um campo magnético oscilatório que, como já foi referido, pode ser detetado por uma bobina recetora de rádio frequências. Esta consiste essencialmente numa ou mais espiras de fio, por onde as linhas de fluxo magnético podem passar. À medida que o campo magnético oscila, é induzida uma pequena tensão alternada na bobina proporcional a M . No equilíbrio, os núcleos precessam com fases aleatórias causando um cancelamento de momentos magnéticos e, por esse motivo, não produzem nenhum sinal detetável. Este comportamento é visível na Figura 2.1[19].

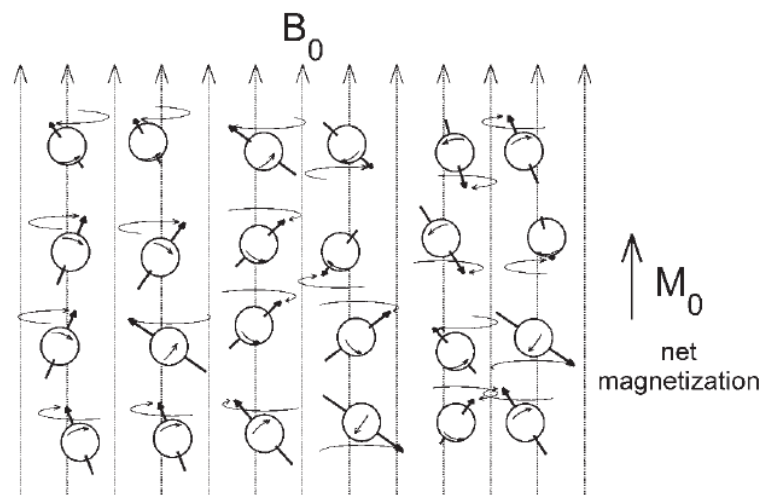


Figura 2.1- Representação de vários núcleos que precessam em torno de um campo magnético, B_0 , com fases aleatórias, não produzindo magnetização transversal. Contudo, ligeiramente mais núcleos estão orientados no mesmo sentido de B_0 , dando origem a uma pequena magnetização longitudinal, M_0 [19].

Embora, no equilíbrio, a magnetização transversal seja nula devido à aleatoriedade de fases, existe uma pequena magnetização longitudinal, M_0 , paralela a B_0 . Esta magnetização existe porque os núcleos têm uma pequena preferência em alinharem-se com o mesmo sentido do campo externo, uma vez que este é um estado de menor energia. Ao valor de M_0 dá-se o nome de magnetização de equilíbrio. Um exemplo desta magnetização está visível na Figura 2.1 [19].

A excitação por rádio frequência é similar ao mecanismo envolvido ao empurrar um balanço de uma criança. Energia pode ser transferida para o balanço empurrando à sua frequência natural, ou de ressonância. À medida que este ganha energia, a amplitude do seu movimento começa a aumentar gradualmente. Quando a força deixa de ser aplicada a frequência de ressonância do balanço é idêntica à frequência que ele oscila sozinho. Da mesma forma, os núcleos da amostra podem ser obrigados a precessar em sincronia aplicando um campo magnético transversalmente a B_0 . O campo aplicado, denominado B_1 , necessita de ter uma frequência de rotação exatamente igual à frequência de Larmor do núcleo alvo para satisfazer a condição de ressonância. Como a frequência de Larmor cai na banda da radiofrequência, este processo é chamado de excitação por rádio frequência. Quando o campo B_1 é desligado, a magnetização transversal continua à frequência de Larmor, produzindo um campo magnético, à custa do campo longitudinal, que pode ser detectado. Eventualmente, forças intermoleculares e internucleares provocam a perda de fase e a magnetização transversa acaba por decair [19], [20].

Quando B_1 deixa de ser aplicado, a perda de energia dos núcleos resulta simultaneamente numa redução do campo magnético transversal e numa recuperação do campo magnético longitudinal para o seu valor de equilíbrio. Estes processos são chamados de relaxação transversal e recuperação longitudinal, respetivamente. Como os tempos destes processos dependem do conjunto molecular sujeito ao campo magnético, estes podem ser usados para diferenciar diferentes tecidos numa imagem de ressonância [19], [20].

À medida que a magnetização transversa decai, a magnetização longitudinal é lentamente restaurada para o seu valor de equilíbrio, M_0 . A variável de tempo relacionada com esta recuperação é chamada de T1 e depende da força do campo magnético B_0 e do meio da amostra [20]. Por outro lado, a variável de tempo relacionada com o decaimento da magnetização transversal é chamada de T2. Esta é mais rápida que T1, pois estão envolvidos alguns processos adicionais [19].

Para produzir uma imagem, cada *slice* de tecido é excitado e o sinal é amostrado como uma função de tempo depois de cada excitação. O sinal recebido é estudado em

frequência usando a transformada de Fourier. A análise pode ser feita a partir de T1 ou T2, dando origem a imagens *T1-weighted*, ou T1w, e *T2-weighted*, ou T2w, respetivamente [19].

Tecidos com grande densidade estão sujeitos a grandes forças de inércia num campo magnético e, por isso, têm tempos de relaxação mais curtos do que tecidos menos densos. Em imagens T1w, amostras com grande T1 têm intensidades baixas, enquanto amostras com baixo T1 têm intensidades elevadas. De forma inversa, imagens T2w têm intensidades elevadas para T2 longos e intensidades baixas para T2 curtos. As imagens T1w são as mais indicadas para distinguir estruturas anatómicas, enquanto as imagens T2w são mais uteis na identificação de patologias [21].

A ressonância magnética é um dos métodos de imagiologia com maior capacidade de contraste entre tecidos e consegue obter imagens em três dimensões. Para além disso, evita o uso de radiação ionizante, prejudicial ao organismo, não se conhecendo nenhum efeito adverso à sua utilização. É, no entanto, um método imagiológico relativamente caro e lento, e é impossibilitado quando os pacientes têm implantes ferromagnéticos [20], [22].

2.3. O encéfalo

Em média, o cérebro humano pesa cerca de 1500 gramas, correspondendo a aproximadamente 2% do peso total de um adulto. É uma espécie de massa gelatinosa revestida por três membranas chamadas de meninges e protegida por uma capsula exterior de osso, o crânio. O encéfalo está suspenso em líquido cefalorraquidiano, suportando-o e absorvendo embates em movimentos rápidos da cabeça. As principais artérias e veias cerebrais estão localizadas nas meninges [23].

2.3.1. Orientação no encéfalo

As relações espaciais do encéfalo e da espinal medula são referentes ao eixo imaginário que passa por estas duas estruturas. Este eixo é chamado de *neuraxis* e assemelha-se a uma forma em T. A parte vertical é a linha que passa por toda a espinal medula e o tronco encefálico, e a parte horizontal é entendida entre o polo frontal e occipital. Este eixo está representado na Figura 2.2 [23], [24].

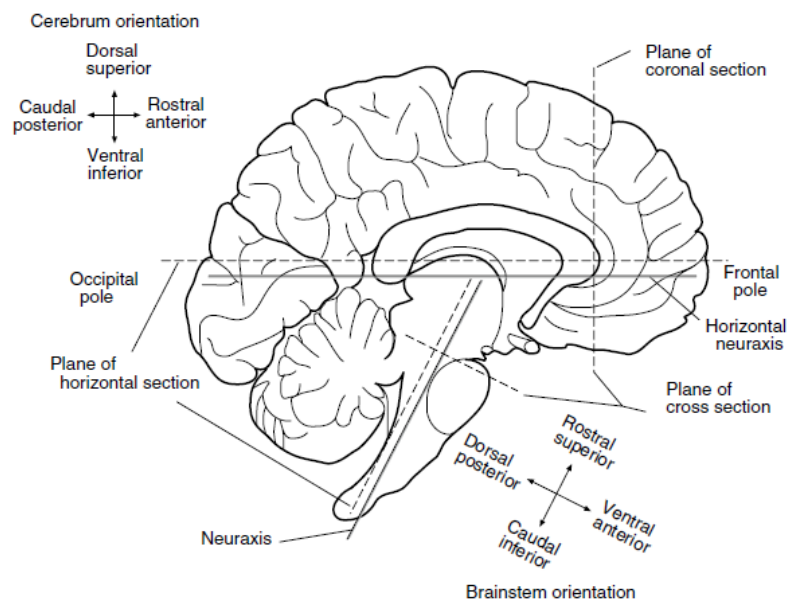


Figura 2.2- Esquemático do plano sagital do encéfalo. Representação do *neuraxis*, do plano axial e coronal e da orientação [24].

O *neuraxis* é usado como ponto de referência na criação dos três diferentes planos direcionais que seccionam o encéfalo: axial, sagital e coronal. O plano axial é paralelo ao *neuraxis*, enquanto o plano coronal é perpendicular ao mesmo. O plano sagital é o plano vertical que passa pela fissura longitudinal. A Tabela 2.1 resume as orientações convencionadas que permitem navegar pelas secções do encéfalo [23], [24].

Tabela 2.1- Descrição das orientações convencionadas para descrever o posicionamento no encéfalo. Os planos indicados correspondem às vistas cujas orientações provocam mudança do slice observado [23], [24]

Plano	Orientações	Descrição
Coronal	Rostral	Também conhecida como anterior. No sentido do polo frontal do encéfalo.
	Caudal	Também conhecido como posterior. Direção oposta à rostral, no sentido ao polo occipital do encéfalo.
Axial	Dorsal	Também conhecida como superior. No sentido da secção do encéfalo mais próxima do topo da cabeça.
	Ventral	Também conhecida como inferior. Direção oposta à dorsal, no sentido da secção do encéfalo mais próximo do pescoço.
Sagital	Medial	No sentido da fissura longitudinal.
	Lateral	Direção oposta à medial, no sentido à fronteira do encéfalo.

2.3.2. Anatomia

O sistema nervoso coordena todas as ações voluntárias e involuntárias, e é responsável por transmitir sinais entre diferentes partes do corpo. Anatomicamente pode ser dividido em sistema nervoso central, composto pelo encéfalo e espinal medula, e sistema nervoso periférico, formado pelos nervos emergentes do encéfalo e da espinal medula [23].

No esquema médico, o encéfalo é normalmente dividido em quatro regiões. Estas são o telencéfalo, o diencéfalo, o tronco encefálico e o cerebelo. O telencéfalo e o diencéfalo compõem os hemisférios cerebrais, e o tronco encefálico é constituído pelo mesencéfalo, a ponte e o bulbo raquidiano. Uma ilustração destas regiões está presente na Figura 2.3 [23], [25].

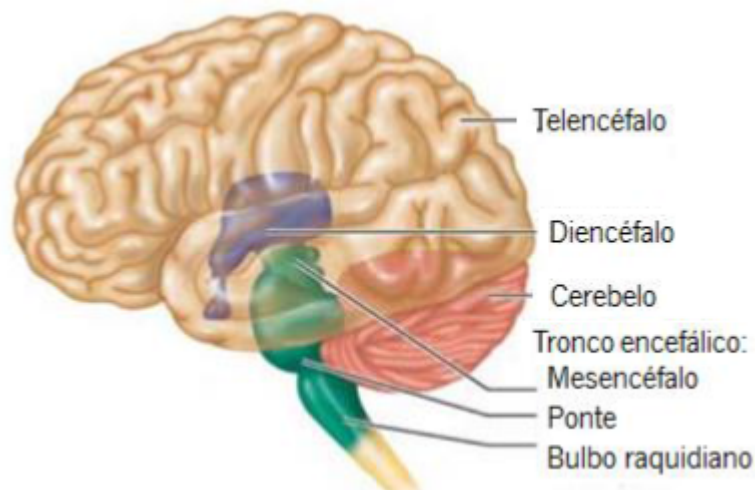


Figura 2.3- Principais subdivisões do sistema nervoso central visto num plano sagital. Adaptado de [25].

O padrão básico do sistema nervoso central é uma cavidade rodeada por substância cinzenta (corpos de células nervosas), existindo externamente substância branca (tratos de axónios mielinizados). A espinal medula obedece a este padrão, mas o telencéfalo tem uma camada exterior adicional de substância cinzenta chamada de córtex. Uma vista em corte do tecido do telencéfalo é visível na Figura 2.4 c). À medida que se desce pelo tronco encefálico o córtex desaparece, mas núcleos dispersos de substância cinzenta, chamados de gânglios da base, são encontrados na substância branca. No final caudal do tronco encefálico o padrão básico é retomado. A substância branca é responsável por estabelecer ligação entre as diversas regiões cerebrais e as restantes estruturas do sistema nervoso central. Os

gânglios da base têm especial importância no controle motor, mas também está associado ao raciocínio e emoções. O córtex cerebral é onde a “consciência” é encontrada. É responsável pelo pensamento, emoções, sensações, comunicação e percepção [25].

As cavidades ocas centrais no interior do encéfalo são chamadas de ventrículos. As camaras ventriculares estão interligadas, preenchidas com líquido cefalorraquidiano e revestidas por células ependemárias [25].

Na superfície, os hemisférios são marcados por incisões chamadas de sulcos. Às saliências entre dois sulcos dá-se o nome de giro. Geralmente, aos sulcos mais profundos e constantes dá-se o nome de fissura. Os sulcos e giros mais predominantes são pontos de referência anatómicos importantes, similares a todos os humanos. Os dois hemisférios são separados pela fissura longitudinal, enquanto outra grande fissura, a fissura cerebral transversal, separa os hemisférios do cerebelo [25].

Vários sulcos dividem cada hemisfério em cinco lobos: frontal, parietal, occipital, temporal e ínsula. Na Figura 2.4 estão representados os principais sulcos e fissuras responsáveis pela divisão dos hemisférios nos seus respectivos lobos [25].

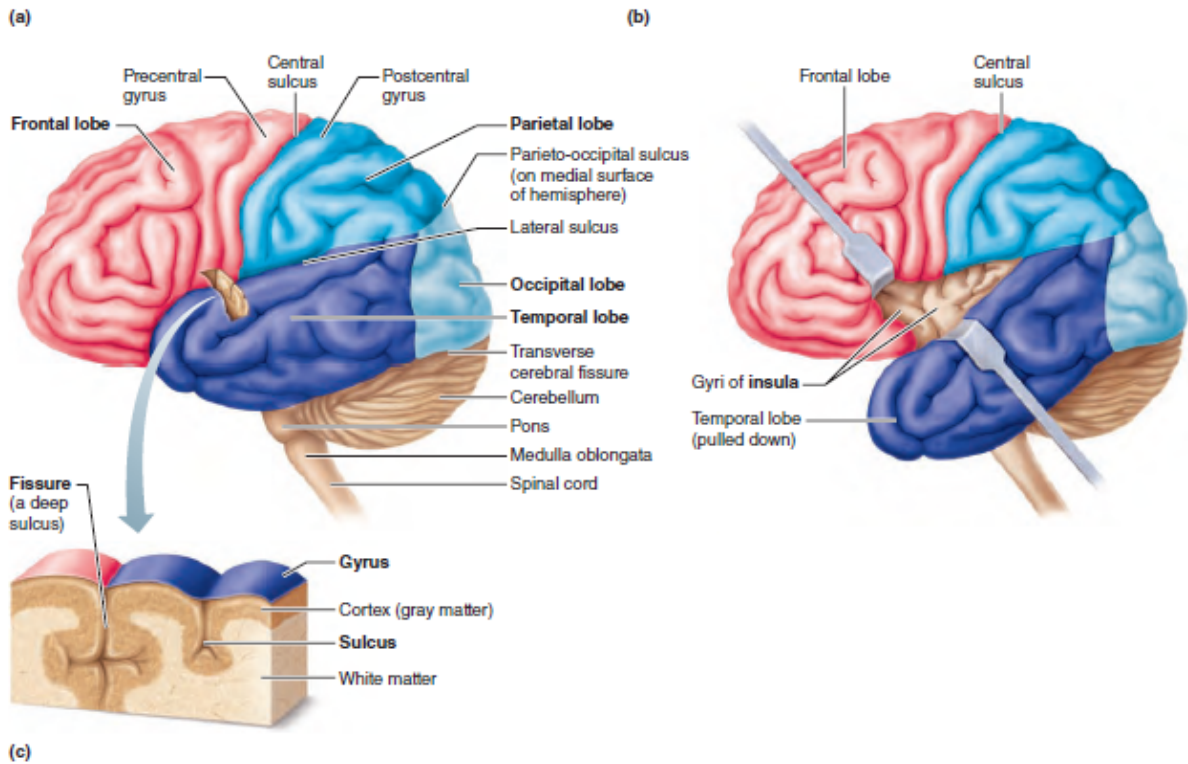


Figura 2.4- Representação dos principais sulcos e fissuras do encéfalo: a) vista lateral; b) localização da ínsula; c) vista em corte do tecido cerebral. Adaptado de [25].

Os lobos cerebrais foram, originalmente, uma classificação puramente anatómica. No entanto, e embora grande parte das atividades requerem coordenação entre várias partes do cérebro, foi comprovado que cada lobo está associados a funções cerebrais específicas [24], [25]. Assim, o lobo frontal é responsável pelo pensamento, movimento voluntário, personalidade e comunicação verbal, sendo a parte mais desenvolvida em relação a outros animais; o lobo parental tem um papel importante na integração sensorial dos vários sentidos; o lobo occipital é responsável pela interpretação de sensações visuais e pelo movimento voluntário dos olhos; no lobo temporal dá-se a interpretação de sinais visuais e é onde se armazenam as memórias; por fim, a ínsula está envolvida na resposta autónoma das vísceras relativamente a sensações como olfato, paladar e dor [23]–[25].

Para além dos lobos, existem ainda redes de neurónios que trabalham em conjunto mas que, por se expandirem ao longo de grandes distâncias, não podem ser localizadas em regiões específicas. Um exemplo é o sistema límbico [25].

O sistema límbico é um grupo de estruturas, localizadas medialmente nos hemisférios e no diencéfalo, que rodeiam a parte superior do tronco encefálico. A maioria

do sistema é responsável por mediar respostas emocionais, mas algumas estruturas, como o hipocampo e a amígdala, estão relacionadas com a memória [25]. O hipocampo é uma das estruturas cerebrais mais estudadas uma vez que é uma das primeiras regiões a ser afetadas por doenças neurodegenerativas, como o caso do Alzheimer [7].

2.3.3. Hipocampo

O hipocampo é um dos componentes principais do encéfalo de todos os vertebrados. Os humanos, tal como todos os mamíferos, têm dois hipocampos, um em cada um dos lobos temporais, arqueados em torno do mesencéfalo [24], [26]. Embora seja uma estrutura com alta variabilidade de sujeito para sujeito, o seu arranjo geral, tal como é visível na Figura 2.5, assemelha-se a um cavalo-marinho [26].

De uma forma geral, admite-se que o hipocampo desempenha um papel fundamental na aprendizagem e na memória. Informação proveniente de grandes zonas corticais converge para a área entorrinal e, posteriormente, para o hipocampo. Assim, a informação recém-adquirida passa pelo “filtro” hipocampal antes de ser fixada no neocórtex, sendo possível distinguir a memória de curto prazo, que está relacionado com hipocampo, com a memória de longo prazo, que depende do neocórtex [26].

O hipocampo está associado a todos os aspetos da memória semântica, que envolve a memória de factos e conceitos, a memória episódica, que permite recordar eventos e relações entre eles, e a memória espacial, que envolve o reconhecimento da localização espacial. Vale apenas ainda referenciar que uma pequena parte do córtex hipocampal (chamado de allocórtex) tem funções cognitivas. É ainda aceite que o hipocampo intervém na regulação de algum comportamento emocional, especialmente aquele produzido por reações à dor [26].

Apesar da extrema importância do hipocampo, muito raramente este é danificado isoladamente. Na maioria dos casos, o hipocampo é sujeito às mesmas patologias que afetam outras áreas corticais, como tumores, malformações vasculares e degeneração cortical. Para além disso, o hipocampo é notoriamente vulnerável a danos devidos a isquemia, hipoxia, trauma e hipoglicemia. Existem ainda instâncias em que o envolvimento da formação hipocampal é crítico na manifestação da doença; notoriamente no Alzheimer e

na epilepsia no lobo temporal, duas das mais prevalentes doenças neurológicas. Estão ainda a ser descobertos vários indícios do envolvimento do hipocampo na esquizofrenia, outra patologia neurológica extremamente comum [7].

Apesar da existência de mais de um século de estudos neuroanatômicos, e de dezenas de milhares de artigos científicos sobre o hipocampo, não existe ainda um consenso sobre alguns aspetos da sua composição. Na presente dissertação, o hipocampo é considerado como uma das regiões cerebrais do sistema funcional chamado formação hipocampal [7]. A restante formação hipocampal é composta pelo subículo, pré-subículo, para-subículo e o córtex entorrinal [7].

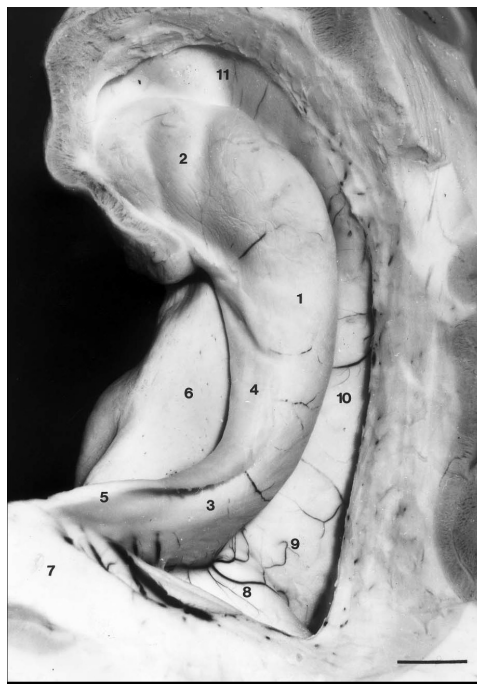


Figura 2.5- Aspeto intraventricular do hipocampo. Aqui são visíveis alguns dos tecidos adjacentes ao hipocampo, como o subículo (6) e a fimbria (4), e os seus três segmentos: corpo (1), cabeça (2) e a cauda (3) [26].

O hipocampo é dividido em três segmentos: corpo, cabeça e cauda [26]. Considera-se ainda que, internamente, o hipocampo é composto por duas estruturas interligadas: o giro dentado e o Corno de Amon (CA). O posicionamento típico destas formações mantém-se constante da cabeça à cauda, existindo a possibilidade de pequenas alterações nas extremidades. Ambas possuem uma organização interna trilaminada, composta por células granulares no giro dentado e células piramidais no CA. O CA é ainda dividido nos setores

CA1, CA2 e CA3 [26]. Cada uma dessas regiões mantém um padrão organizado de conexões intrínsecas e extrínsecas, sendo que a principal aferência para o hipocampo origina-se no córtex entorrinal. As fibras que deixam o córtex entorrinal em direção ao hipocampo constituem a via perfurante e inervam os dendritos das células granulares do giro denteado. Os axônios das células granulares (fibras “musgosas”) projetam-se para as células piramidais da região de CA3, que por sua vez emitem fibras para a região de CA1. De CA1, as fibras projetam-se para o complexo subicular e então para as camadas profundas do córtex entorrinal (feixe angular). Este circuito é normalmente chamado de via tri-sináptica e tem especial influência na memória. Todas as ligações e a organização da formação hipocampal estão ilustradas na Figura 2.6 [7].

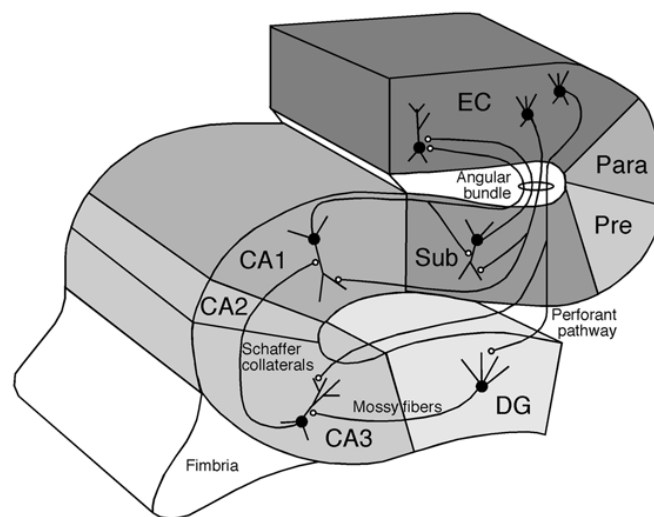


Figura 2.6- Ilustração da estrutura e das principais ligações da formação hipocampal: hipocampo (três zonas CA3, CA2 e CA1), giro denteado (DG), subículo (Sub), pré-subículo (Pre), para-subículo (Para) e córtex entorrinal (EC) [7].

No caso de certas patologias neurológicas, existe uma forte correlação do atrofiamento hipocampal com o avanço da doença. Assim, uma análise volumétrica do hipocampo permite diagnosticar ou acompanhar a progressão de doenças neurológicas. A imagiologia médica, como a ressonância magnética, surge como uma boa ferramenta uma vez que, ao contrário da histologia, permite estudar o hipocampo *in vivo* [27]. No entanto, em contraste com a classificação de tecido cerebral, onde a intensidade do sinal da ressonância magnética pode ser usado para classificar diferentes tipos de tecido,

segmentações anatómicas são mais complexas, uma vez que uma estrutura pode ser composta por vários tipos de tecido [28].

A análise volumétrica do hipocampo, e das suas estruturas vizinhas, em MRI já foi alvo de intensa investigação, tendo sido realizados vários estudos em adultos saudáveis, de diferentes faixas etárias, e em pacientes com uma variedade de problemas neurológicas e psiquiátricos [27].

Um problema recorrente é a variabilidade entre protocolos de medida das estruturas. Embora existam muitas similaridades, os investigadores diferem, principalmente, nas medidas das extremidades, onde alguns protocolos tentam contabilizar qualquer hipocampo visível, enquanto outros optam por maiores níveis de confiabilidade, obtidos por começar e parar em localizações que não cobrem toda a extensão hipocampal. Embora as diferenças possam parecer pequenas, acabam por contribuir para a variabilidade dos resultados entre diferentes laboratórios. Mais importante ainda, é a forma como os laboratórios lidam com as diferenças das dimensões das cabeças dos pacientes, e se reportam os resultados com volumes reais ou normalizados [27].

Uma vez que um especialista experiente demora até a duas horas a segmentar os dois hipocampos de um paciente numa imagem de MRI, a segmentação manual é considerada demasiado ineficiente para o uso clínico de rotina [27]. Felizmente, em conjunto com os avanços nas técnicas de aquisição de imagem, tem havido um aumento no desenvolvimento de técnicas analíticas capazes de estudar alterações no hipocampo. Embora com custos computacionais elevados, vários dos métodos propostos são completamente automáticos, capazes de classificar estruturas específicas da formação hipocampal. No entanto, a maioria dos algoritmos apresenta dificuldades pontuais, especialmente em sujeitos com anormalidades anatómicas, pelo que requerem que indivíduos especializados controlem os resultados, de forma a assegurar níveis de exatidão aceitáveis [27], [29]. Consequentemente, o desenvolvimento de técnicas de segmentação precisas, robustas e fiáveis tem-se tornado um importante desafio na análise quantitativa de MRIs [28].

Avanços de metodologia na volumetria hipocampal automática têm permitido vários laboratórios estudar dimensões de centenas de sujeitos, fornecendo novos dados sobre o comportamento hipocampal durante o seu desenvolvimento, envelhecimento ou quando

sujeito a doenças neurodegenerativas. As observações confirmaram que, embora uma perda da área cortical seja notória devido ao envelhecimento normal, existe uma redução mais acentuada devido ao Alzheimer ou outras doenças neurodegenerativas [27].

2.4. Sumário

O Alzheimer é o tipo mais comum de demência, sendo, na atualidade, a principal doença responsável pela deterioração progressiva do cérebro. Embora muito dos mecanismos, causas e fatores de risco sejam desconhecidos, é estimado que os danos no cérebro começam, entre 10 a 20 anos, antes que qualquer sintoma seja evidente. Nenhuma cura é ainda conhecida, mas um tratamento precoce ajuda a atenuar os sintomas doença. O diagnóstico clínico da doença é um processo longo e penoso para o paciente. Para além disso, existe sempre um certo grau de subjetividade e a eficácia do diagnóstico pressupõe um certo grau de avanço da patologia. A imagiologia por ressonância magnética fornece uma forma não invasiva de detetar, antes do surgimento de qualquer sintoma, a atrofia de certas estruturas cerebrais.

O encéfalo é o órgão central do sistema nervoso, sendo responsável pela coordenação de todas as ações voluntárias e involuntárias. Nos vertebrados, um dos seus principais componentes é o hipocampo. Os humanos, tal como todos os mamíferos, têm dois hipocampos, um em cada um dos lobos temporais, arqueados em torno do mesencéfalo. De uma forma geral, admite-se que o hipocampo desempenha um papel fundamental na aprendizagem e na memória, sendo uma das primeiras estruturas a sofrer neurodegeneração devido ao Alzheimer.

Capítulo 3 Estado da arte

Uma quantidade substancial de literatura sugere que anormalidades no LCR, A β , T-tau e PET precedem neurodegeneração no Alzheimer. No entanto, enquanto o grau de variação deste biomarcadores pode abrandar antes do aparecimento de atrofia estruturais, a neurodegeneração visível por imagiologia aumenta à medida que os sintomas clínicos aparecem. Assim, estudos volumétricos, ou medidas baseadas em vóxeis da atrofia cerebral, retêm uma relação aproximada com o declínio da performance cognitiva em pacientes com Alzheimer [30].

Tanto o MRI como a tomografia axial computadorizada (TAC) foram usados como forma de obter informação estrutural de tecido atrofiado devido a Alzheimer. Contudo, o MRI tem várias vantagens comparado com a TAC: maior resolução, excelente discriminação de matéria cinzenta e branca e capacidade de identificação de pequenas lesões vasculares. Estes fatores contribuem para uma maior sensibilidade e especificidade no diagnóstico de Alzheimer [11].

Alguns dos primeiros estudos efetuados com MRI em pacientes com Alzheimer focaram-se na medição da atrofia do hipocampo, devido à importância funcional desta estrutura na memória. Os resultados verificaram uma redução do volume entre 30% a 40%, comparado com os controlos, dependendo do protocolo de segmentação manual usado e do grau da doença. Da mesma forma, a exatidão na diferenciação entre pacientes com Alzheimer e os controlos oscilava entre 85% e 100% [31].

Autópsias de pacientes, em fases iniciais da doença, mostraram o aparecimento de nós neurofibrilares no córtex entorrinal antes de qualquer envolvimento do hipocampo. Esta descoberta fez com que vários grupos de investigação direcionassem a sua pesquisa para outras estruturas do lobo temporal medial como a amígdala, o córtice parahipocampal, o giro parahipocampal e, principalmente, o córtex entorrinal. Em vários dos estudos, o grau de atrofia do córtex entorrinal foi superior ao do hipocampo, conferindo-lhe um maior poder de separação entre pacientes com Alzheimer e controlo. Contudo, as reduzidas dimensões e alta variabilidade de forma tornam a identificação desta estrutura difícil em imagens de RM, levando a uma maior variabilidade nos resultados [30], [32].

De uma forma geral, melhores segmentações (e conseqüentemente, melhores biomarcadores de Alzheimer) são obtidas em métodos de delimitação do hipocampo, pelo que esta foi a estrutura focada nesta Dissertação. Neste capítulo algumas das dificuldades e desvantagens da segmentação manual de estruturas cerebrais são apresentadas. Posteriormente, é feita uma descrição de alguns dos algoritmos mais importantes na segmentação do hipocampo.

3.1. Protocolos de segmentação manual do hipocampo e atlas cerebrais

Em imagem médica, o objetivo da segmentação é atribuir *labels* aos vóxeis, de tal forma que as regiões de diferentes unidades estruturais ou funcionais sejam classificadas e possam providenciar resultados interpretáveis. A qualidade da segmentação afeta diretamente a visualização, exatidão da subsequente análise e o planejamento da intervenção [33].

Dependendo da aplicação e da modalidade de imagem, um grande leque de métodos de segmentação têm sido propostos de forma a segmentar o hipocampo. No que diz respeito ao envolvimento humano, estes podem ser classificados como: segmentação manual, segmentação semiautomática ou segmentação automática [33].

A segmentação manual é normalmente efetuada por especialistas treinados que, seguindo um protocolo específico, definem as regiões de interesse da imagem. O conhecimento anatómico do especialista contribui para a segmentação quando as fronteiras das estruturas são difíceis de detetar [33].

Embora o hipocampo seja uma estrutura complexa situada no interior do lobo temporal medial, a maioria das suas fronteiras são visíveis com a alta resolução das imagens de RM. De forma a normalizar o delineamento manual em diferentes imagens, diferentes protocolos de segmentação foram sugeridos, capazes de traduzir definições anatómicas para o domínio da imagem. [33].

As diferentes *guidelines* anatómicas foram usadas por vários estudos volumétricos do hipocampo, o que contribuiu para as variações nos resultados reportados. Em 2009, Konrad

et al. avaliou 71 protocolos publicados na literatura, identificando cinco áreas principais onde os protocolos diferem [34]:

- inclusão/exclusão da matéria branca do álveo e fimbria;
- fronteira anterior do hipocampo-amígdala;
- fronteira posterior para onde a cauda do hipocampo se estende;
- fronteira medial interior;
- uso de linhas arbitrárias.

As segmentações manuais permitiram a criação de atlas, isto é, imagens classificadas e validadas para definir as fronteiras de áreas anatómicas ou funcionais de diferentes imagens [33]. Um dos primeiros atlas largamente usados em estudos neuroimagiológicos foi sugerido por Talairach e Tournoux, em 1998. Este foi criado a partir de um cérebro dissecado [35]. Mais tarde, o Montreal Neurological Institute (MNI) propôs um novo atlas baseado em 305 imagens de RM registadas linearmente, ou seja, transformadas para o mesmo espaço físico. O atlas MNI305 foi criado através da média das intensidades de todas as imagens [36], [37]. Em 2001, o Atlas ICBM152 foi criado a partir de 152 imagens adquiridas com uma maior resolução e contraste do que as usadas no MNI305 [38].

Embora a segmentação manual continue a ser o padrão de qualidade no que toca a segmentação de estruturas em MRI, na prática, uma abordagem manual depende fortemente do protocolo usado, sofre de baixa reprodutibilidade e grande variabilidade intra e inter técnico, resultando em grandes diferenças nos resultados finais [30], [33]. Além disso, estudos de grande escala são impraticáveis uma vez que, tipicamente, o hipocampo demora até 2 horas a ser segmentado à mão; um trabalho altamente especializado, tedioso e fatigante. Assim, várias técnicas automáticas foram sugeridas de forma a reduzir ou eliminar a necessidade do processo manual, uma fonte de variabilidade entre segmentações da mesma imagem [39].

3.2. Algoritmos de segmentação do hipocampo

Alguns dos primeiros métodos de segmentação computadorizada seguiram uma abordagem semiautomática onde, de forma a reduzir o trabalho, o utilizador é obrigado a

identificar um conjunto disperso de pontos de referência na imagem, de forma a restringir a subsequente segmentação automática [39]. Embora reduza consideravelmente o tempo de segmentação, este tipo de métodos continua a depender fortemente do seu processo de inicialização e, para estudos de grande escala, delimitação manual pode não ser viável. Como resultado, muita da pesquisa tem sido direcionada para a segmentação completamente automática [39]. Os algoritmos de segmentação do hipocampo são maioritariamente baseados em crescimento regional ou modelos deformáveis, métodos baseados em aparência e forma ou métodos baseados em atlas [33].

3.2.1. Crescimento regional ou modelos deformáveis

Algoritmos de crescimento regional ou de modelos deformáveis começam, normalmente, de um conjunto inicial de pontos de referência, que se deforma ou expande de modo a coincidir com a estrutura a ser segmentada. Algumas das primeiras abordagens funcionavam de forma semiautomática, sendo a inicialização feita por um especialista. Em Ashton et al. (1997), a região começa com uma linha de pontos ao longo do eixo do hipocampo e deforma-se elasticamente, sendo constringido por tensões de superfície, desvios da superfície esperada e forças de tecidos vizinhos [40].

Em 1998, Ghanei aplicou modelos de contorno dinâmicos, introduzidos inicialmente por Terzopoulos et al., à segmentação do hipocampo. A geometria do modelo deforma-se devido a forças internas, baseadas curvatura local do modelo, e forças externas, calculadas a partir de características da imagem obtidas por métodos de deteção de bordas. Embora capaz de bons resultados de segmentação, o método depende fortemente do método de inicialização [41].

O algoritmo SACHA (*Segmentation Automatisée Compétitive de l'Hippocampe et de l'Amygdale*), desenvolvido por Chupin et al. em 2007, segmenta simultaneamente o hipocampo e a amígdala. É inicializado por uma definição manual de uma caixa delimitadora e de pontos de referência, colocados pelo operador em cada uma das estruturas. A deformação da região é regularizada por *Markov random fields* (MRF) [42], [43].

Com uma definição automática de um ponto de referência, o FMASH (*fast marching for automated segmentation of the hippocampus*), por Bishop et al. em 2011, propaga a

região ao longo do caminho com a menor resistência, definido por uma função de potencial da intensidade da imagem usando o método *fast marching*, em 3D [44], [45].

3.2.2. Métodos baseados em aparência e forma

Os modelos de forma ativa foram usados em segmentação de imagem médica por Shen et al. em 2002. Através do ajuste de um modelo de forma paramétrico, o algoritmo usado é capaz de detectar automaticamente características da imagem ou certos pontos de referência. Shen usou o modelo de forma introduzido por Cootes et al. onde, de modo a aumentar a especificidade da segmentação, a deformação do modelo é restringida de acordo com a informação presente num conjunto de treino. Embora o método apresente resultados estatisticamente equivalentes a diferenças inter-avaliador para a segmentação manual, o algoritmo recorre a um número de pontos de fronteira que ajudam a guiar a deformação do modelo para as fronteiras corretas, quando estas não são claramente definidas na imagens de ressonância magnética [46], [47].

O modelo de aparência ativo, introduzido originalmente por Cootes et al. em 2001, para além de ter em conta a forma das estruturas a segmentar, considera informação proveniente da intensidade e textura da imagem, ou seja, a “aparência”. Em 2007, Babalola et al. usou o perfil de intensidades ao longo da normal da fronteira da estrutura, o modelo de aparência de perfil. A segmentação é produzida correspondendo o modelo com a imagem e minimizando o quadrado das diferenças [48]. Em 2011, Patenaude et al. incorporou este tipo de modelo numa *framework* bayesiana, permitindo relações probabilísticas entre formas e intensidades ser completamente exploradas [49].

Hu e Collins propuseram, em 2007, um algoritmo de segmentação completamente automático baseado num modelo que combina métodos de superfície de nível, para modelar a forma das estruturas cerebrais, e as suas variações com modelos de aparência ativa, para gerar imagem que são usadas na segmentação. O algoritmo RLS (*recursive least square*) é adotado para minimizar a diferença entre as imagens de teste e a imagem sintetizada. Este método usa imagens de várias modalidades para melhor a performance da segmentação [50].

3.2.3. Métodos baseados em atlas

Atlas cerebrais são definidos por *labels* anatômicos num espaço estereotático, isto é, um sistema de coordenadas normalizado que estabelece uma transformação de um vóxel num cérebro para o vóxel correspondente num segundo cérebro. Alinhando uma nova imagem com *labels* pré-definidos num espaço de referência, permite usar conhecimento prévio para processar uma imagem não segmentada. Uma distinção pode ser feita entre atlas probabilísticos, que atribuem a cada vóxel uma probabilidade de observar uma estrutura particular, e atlas que correspondem à segmentação manual de uma imagem de um cérebro individual [29].

A transformação de um atlas manualmente classificado para o sistema de coordenadas de uma imagem em estudo permite o uso do mapa de *labels* resultante, para obter uma segmentação no espaço alvo. Entre os primeiros trabalhos de segmentação cerebral baseada em atlas, estão os publicados por Collins et al. e Christensen et al. em meados dos anos 90. Nestes trabalhos, um único atlas é alinhado, não linearmente, com a imagem alvo, e a transformação resultante é usada para propagar os mapas de *labels* para a imagem em estudo. A Figura 3.1 ilustra o conceito de segmentação baseada em atlas [29]. Alternativamente, métodos baseados em modelos de aparência também foram propostos para guiar o registo, como os sugeridos por Duchesne et al. em 2002 e por Klemencic et al. em 2004.

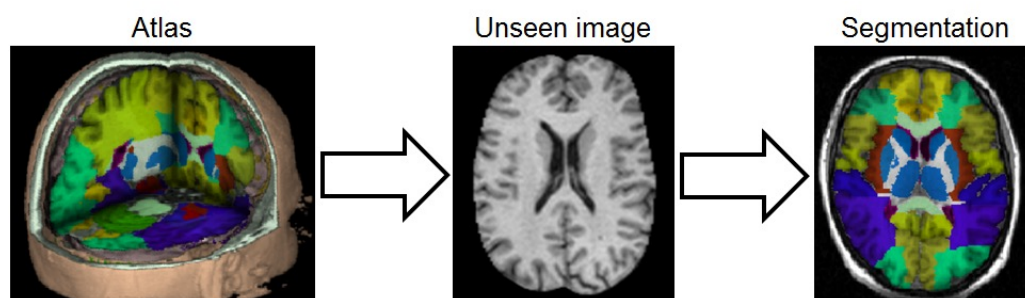


Figura 3.1- Segmentação cerebral baseada em atlas: um atlas é registrado com uma nova imagem e os *labels* são propagados [29].

Esta abordagem depende crucialmente do alinhamento do atlas com a imagem alvo. Por consequência, a segmentação resultante falha em áreas onde o registo subjacente falha [29].

Diferentes métodos de registo elásticos (Miller et al.,1993; Losifescu et al.,1997; Crum et al.,2001), fluidos (Haller et al.,1996; Hogan et al., 2000; Crum et al.,2001) e baseados em fluxo ótico (Dawant et al., 1999; Hartmann et al., 1999; Baillard et al.,2001) já foram usados para corresponder uma imagem de um paciente a um atlas [33].

No algoritmo ANIMAL (*Automatic Nonlinear Image Matching and Anatomical Labeling*), proposto por Collins et al. em 1995, a imagem do paciente em estudo é usada como ponto de partida, e o atlas como o alvo, num registo não rígido (permite deformações locais na imagem alvo para permitir um melhor alinhamento com a imagem de referência). O mapeamento dos *labels* é conseguido por transformação inversa, recuperada pelo ANIMAL [51].

Uma forma encontrada de aumentar a exatidão da segmentação foi o uso de múltiplos atlas, reduzindo assim, a tendência para um único atlas e tomando vantagem da informação presente em vários sujeitos [52], [53]. Vários atlas podem ser transformados para o espaço de coordenadas da imagem alvo, que será segmentada por um registo não rígido. A classificação de cada vóxel da imagem é então determinada por fusão das segmentações deformadas de cada um dos atlas, normalmente com uma regra de votação. Desta forma, erros cometidos por atlas individuais podem ser compensados por outros atlas. Esta abordagem está ilustrada na Figura 3.2. Foi mostrado que esta é uma das melhores abordagens para a delimitação de estruturas subcorticais em MRI [54].

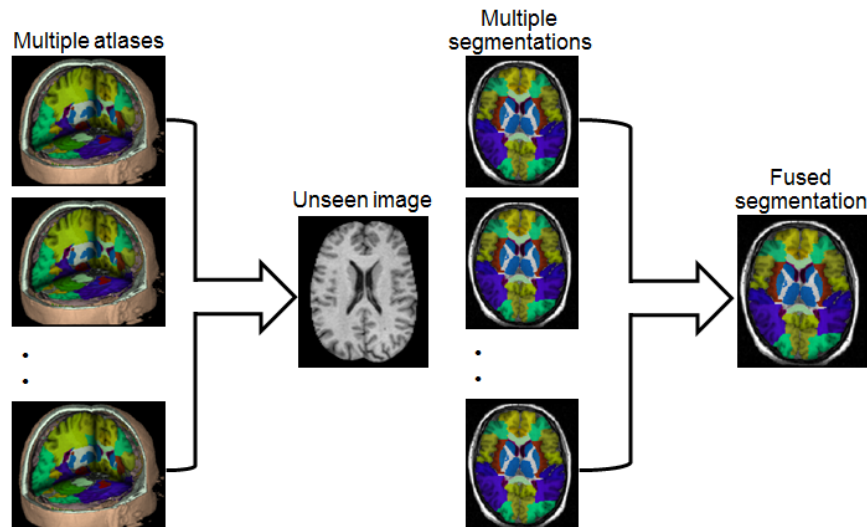


Figura 3.2- Segmentação baseada a múltiplos atlas. As imagens dos atlas são registradas em novas imagem, permitindo a propagação dos seus labels. A segmentação final é obtida por fusão das segmentações individuais [29].

A variabilidade entre sujeitos pode exceder o espaço de transformação procurado pelos algoritmos não rígidos, pelo que erros de correspondência podem ser propagados para o resultado da segmentação. Na prática, a exatidão deste tipo de métodos depende de quatro fatores: o conjunto inicial de atlas, o algoritmo de registo, a estratégia de seleção dos atlas depois do registo e o método de fusão de *labels* [39], [48].

Um conjunto representativo, mas imparcial, de atlas deve ser escolhido para a população de interesse, uma vez que este tende a melhor a performance comparativamente com único e normalizado coorte de atlas [39]. Embora nem sempre esteja disponível, um grande conjunto de atlas ajuda a lidar com a variabilidade em populações muito diversas ou com uma subpopulação específica [33], [55].

No contexto do registo, a abordagem não rígida não assegura a correspondência completa entre as imagens mas, segundo a comparação feita por Carmichael et al., produz melhores resultados do que registo com transformações lineares, especialmente no caso de estruturas com forma complexa, como o hipocampo [39].

Em 2009, Yassa e Stark avaliaram a eficácia de 12 métodos de registo no alinhamento de estruturas do lobo temporal medial, incluindo alinhamento Talairach e 3dWarpDrive no *software* AFNI (Cox, 1996), normalização em SPM (Asburner e Friston, 1999,2005), LDDMM (*Large Deformation Diffeomorphic Metric Mapping*, Beg et al.,2005; Miller et al., 2005),

DARTEL (*Diffeomorphic Anatomical Registration Through Exponential Lie Algebra*, Ashburner, 2007) e o Demons, um algoritmo baseado numa analogia com conceitos da termodinâmica (Thirion, 1998). O algoritmo Demons teve a melhor performance no alinhamento do hipocampo [56].

O problema da seleção de vários atlas foi discutido por Aljabar et al., em 2007, onde propõem uma estratégia de pré-seleção dos atlas, antes de qualquer segmentação. Baseado em métricas de intensidade ou informação do sujeito, todos os atlas são ordenados de acordo com a sua conveniência para a imagem em estudo. A sobreposição do resultado da segmentação obtida e a segmentação real é máxima quando, entre 10 a 20 dos primeiros atlas são selecionados, e combinados segundo uma regra de votação. Quando atlas menos similares são fundidos na segmentação, a informação relevante assume um menor peso, e a eficácia da segmentação converge para um valor inferior [29]

Relativamente à fusão dos *labels*, o método básico consiste na combinação dos mapas de *labels* transformados com uma votação voxel-a-voxel por maioria [53]. Uma abordagem mais sofisticada é a atribuição de pesos aos votos de cada atlas baseados na sua performance, como por exemplo o algoritmo STAPLE (*Simultaneous truth and Performance Level Estimation*). O STAPLE produz uma estimativa probabilística da segmentação através da atribuição de um peso a um número de segmentações individuais, enquanto considera a performance de cada segmentação. Um algoritmo EM (*expectation-maximization*) é usado para iterativamente estimar a segmentação real, pesando todas as segmentações individuais [29], [57].

Em contraste com o STAPLE, onde os pesos são avaliados globalmente no atlas, pesos calculados localmente foram desenvolvidos, selecionando regiões ou voxéis próximos à estrutura em estudo [58]. Em 2009, Artachevarria et al. utilizou a informação da intensidade da imagem alvo e dos atlas, de forma a pesar cada atlas na votação. A votação depende da similaridade local de cada atlas com a imagem alvo [59]. Em 2012, Coupé et al. propôs um método capaz de segmentar e classificar padrões de atrofiamento do hipocampo e do córtex entorrinal, usando uma abordagem de patches não locais. O seu método obteve um dos melhores resultados na separação entre controlos e sujeitos com Alzheimer [60]. De

uma forma geral, a atribuição local de pesos obtém maior exatidão do que quando se considera informação presente em toda a imagem [61].

Métodos baseados em aprendizagem estatística também foram introduzidos como forma de pesagem na segmentação baseadas em múltiplos atlas. Em 2010, Sdika et al. usou um mapa de exatidão de segmentação, para estimar a exatidão média de cada vóxel dos atlas no conjunto de treino, e a estimativa é usada como peso na fusão dos *labels* [62].

O uso de atlas também foi combinado com outras abordagens de segmentação. Em 2008, Lijn et al. usou vários atlas para gerar um mapa de probabilidades, que é depois combinado com um modelo de intensidades e vizinhança para criar uma função de energia. As estruturas são então segmentadas, minimizando a função de energia usando o algoritmo de *graph cuts* [63]. Em 2009, Wolz et al. melhorou o modelo de intensidades e, em 2010, Lötjönen et al. propoz uma abordagem EM para otimizar a função de energia [64], [65].

3.3. Abordagem seguida nesta dissertação

O uso de diferentes métricas de avaliação, de conjuntos de imagem ou de diferentes esquemas na validação fazem com que, a comparação da eficácia entre os diferentes métodos apresentados nem sempre seja possível. Além disso, como a segmentação se destina a ser usada num sistema de *Computed Aided Diagnostic*, o método usado necessita de ser capaz de segmentar hipocampos com elevado grau de atrofia.

Nesta dissertação será usado o método de segmentação baseado em *patches* não locais, proposto por Coupé et al., em 2011, como a base para a criação de uma ferramenta de classificação dos diferentes estados de demência. Este foi escolhido devido à qualidade dos resultados obtidos e à relativa simplicidade do método [28]. Além disso, o algoritmo é expandido, no ano seguinte, de forma a calcular mapas de gradações das estruturas cerebrais, que podem ser usados como um novo biomarcador de Alzheimer [66]. Este sistema foi validado com as imagens da base de dados ADNI, um dos maiores estudos longitudinais de pacientes de Alzheimer, tendo sido reportado resultados de classificação que, ainda hoje, se mantêm competitivos [60].

Como nenhuma versão do algoritmo de Coupé et al. (*Scoring by Nonlocal Patch Estimator*- SNIPE) se encontra disponível publicamente, uma das principais tarefas da dissertação foi a implementação da totalidade do método. É ainda proposta uma alteração ao sistema sugerido por Coupé, sendo a informação das bordas das imagens, obtida pelo filtro Laplaciano, usada para melhorar o cálculo da segmentação e da gradação.

3.4. Sumário

Alguns dos primeiros estudos efetuados com MRI, em pacientes de Alzheimer, focaram-se na medição da atrofia do hipocampo, devido à importância funcional desta estrutura na memória. Estes verificaram reduções significativas do volume da estrutura em pacientes com a patologia. Autópsias de pacientes, em fases iniciais da doença, mostraram o aparecimento de nós neurofibrilares no córtex entorrinal antes de qualquer envolvimento do hipocampo. Contudo, as reduzidas dimensões e alta variabilidade de forma tornam a identificação desta estrutura difícil em imagens de RM, levando a uma maior variabilidade nos resultados.

Embora a segmentação manual seja o padrão de qualidade no que toca a segmentação de estruturas, a delimitação do hipocampo é um trabalho altamente especializado, tedioso e demorado. Assim, várias técnicas automáticas foram sugeridas de forma a reduzir ou eliminar a necessidade do processo manual. Os algoritmos de segmentação do hipocampo são maioritariamente baseados em crescimento regional ou modelos deformáveis, métodos baseados em aparência e forma ou métodos baseados em atlas.

Algoritmos de crescimento regional ou de modelos deformáveis começam, normalmente, de um conjunto inicial de pontos de referência, que se deforma ou expande de modo a coincidir com a estrutura a ser segmentada. Os métodos baseados em aparência e forma modelam a forma da estrutura pretendida e usam o modelo resultante para realizar a segmentação. Os algoritmos baseados em atlas usam imagens com *labels* pré-definidos num espaço de referência, permitindo o uso do conhecimento prévio para processar imagens não segmentadas.

Nesta dissertação será usado o método proposto por Coupé et al., em 2012, como a base para a criação de uma ferramenta de classificação dos diferentes estados de demência

Capítulo 4 Segmentação do hipocampo baseada em patches não locais

Os métodos baseados em patches são bastante competitivos em determinados subdomínios do processamento de imagem, como a síntese de texturas (Efros e Freeman, 2001), super-resolução (Protter et al., 2009), reconstrução (Criminisi et al., 2004) e restauração de imagens (Buades et al., 2005). Em cada uma destas aplicações, a abordagem por *patches* exibe uma performance muito elevada, apesar da sua simplicidade [67]–[70].

Inspirado no filtro de média não local para ruído gaussiano, Coupé et al. propuseram, em 2011, um método baseado em *patches* não locais capaz de segmentar o hipocampo e o ventrículo cerebral em imagens de MR, a partir de segmentações manuais de especialistas. Duas das principais propriedades do filtro de média não local são exploradas para melhorar a segmentação: a redundância natural da informação contida na imagem, que pode ser usada para aumentar drasticamente o número de amostras, e o contexto de intensidade local, que é usado como uma forma robusta de comparação de amostras. O método foi usado para segmentar o hipocampo de sujeitos saudáveis e ventrículo lateral de pacientes com Alzheimer [28].

No presente capítulo será introduzido o conceito de *patch* e de processamento não local. O método de segmentação proposto por Coupé et al. é apresentado, incluindo os algoritmos de pré-processamento e a sua necessidade, e estratégias de aceleração do método.

4.1. Visão geral do método

Tal como nos métodos baseados em atlas, o algoritmo proposto por Coupé et al. usa imagens de RM segmentadas manualmente por especialistas como antecedentes, de forma a segmentar estruturas anatómicas. No entanto, este apresenta duas diferenças principais: a escala dos objetos considerados e o esquema de fusão de *labels* [28], [71].

Enquanto os algoritmos baseados em atlas trabalham ao nível da estrutura anatómica, este método usa uma escala menor através de patches, isto é, pequenos

conjuntos de vóxeis vizinhos. Assim, enquanto a segmentação por atlas realiza uma fusão de estruturas dos templates deformadas não linearmente, a segmentação baseada em patches não locais alcança a classificação de cada vóxel individualmente, por comparação do seu patch envolvente com patches dos sujeitos de treino, onde o label do vóxel central é conhecido. Quando um patch sob estudo se assemelha a um patch dos sujeitos de treino, os seus vóxeis centrais são considerados pertencentes à mesma estrutura. Com este método, várias amostras de cada sujeito de treino podem ser usadas durante a fusão de labels, permitindo um drástico aumento do número de amostras de patches envolvidas no processo [28].

Os algoritmos baseados em atlas usam, normalmente, um esquema de votação por maioria para a fusão de *labels* que considera a relevância, ou peso, de todas as amostras classificadas como semelhantes [53], [72]–[74]. Em contraste, o estimador não local usa a diferença baseada na intensidade entre os *patches* sob estudo para efetuar uma fusão ponderada dos *labels*. O termo “não local” indica que a distância espacial entre os centros dos *patches* não é tida em conta na ponderação; assim, o peso de cada amostra é apenas influenciado pela similaridade de intensidades entre *patches*. Neste tipo de abordagem, a distância baseada na intensidade entre *patches* diminui à medida que a relevância entre as amostras considerada aumenta [28].

Desta forma, o esquema de médias de *patches* não locais permite o uso robusto de um grande número de amostras durante a estimativa, aproveitando a redundância de informação presente na imagem. Este número será mais importante que o número de sujeitos de treino, ao contrário dos métodos baseados em atlas, onde o número de atlas utilizado prevalece. Além disso, contrariamente à votação por maioria, a média não local permite uma robusta distinção das amostras mais similares de acordo com o seu contexto local [28].

4.2. Pré-processamento

O primeiro passo da segmentação por *patches* não locais envolve a organização dos sujeitos de treino para serem usados na comparação de *patches*. Durante esta fase, a variabilidade causada pela formação da imagem é minimizada através de filtragem, correção

das não-homogeneidades e normalização das intensidades. Além disso, uma vez que regularidade anatômica entre os sujeitos será usada para guiar a procura dentro da biblioteca, os sujeitos da base de dados são linearmente transformados para o espaço estereotático de forma a assegurar a correspondência entre as localizações anatômicas das diferentes imagens. A sequência destes passos é visível na Figura 4.1 [28].

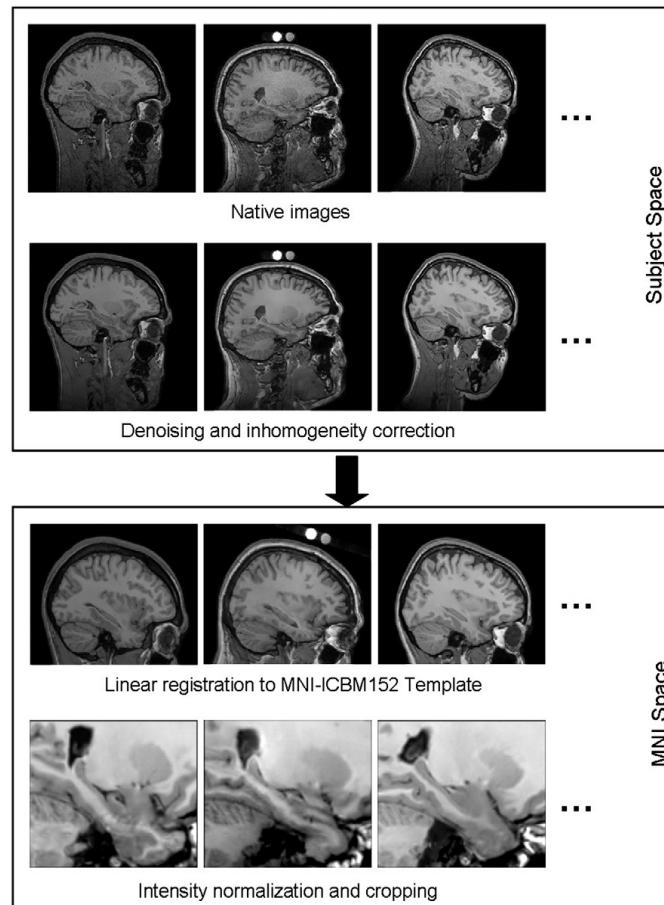


Figura 4.1- Pipeline de pré-processamento usado para a construção da biblioteca de imagens. Primeiro, filtragem e correção das não-homogeneidades são feitas no sistema de coordenadas do sujeito. Então, as imagens são registradas para o espaço do atlas MNI-ICBM152. Finalmente, a normalização das intensidades de cada sujeito é feita antes de cortar as imagens à volta das estruturas de interesse [28].

4.2.1. Filtragem

Todas as imagens da base de dados foram, primeiramente, filtradas com o filtro de média não local 3D proposto por Coupé et al., em 2008. De forma a reduzir o ruído Rician, a adaptação do filtro não local de Wiest-Darselle et al. também foi usada [28].

Na formulação clássica do filtro de média não local, a intensidade restaurada $NL(u)(x_i)$ do voxel x_i é a média pesada de todas as intensidades dos voxéis da imagem u definida como:

$$NL(u)x_i = \sum_{x_j \in \Omega^3} w(x_i, x_j)u(x_j) \quad (4.1)$$

onde $u(x_j)$ é a intensidade do voxel, x_j e $w(x_i, x_j)$ é o peso atribuído a $u(x_j)$ na restauração do voxel x_i . O peso quantifica a similaridade das vizinhanças locais N_i e N_j , dos voxéis x_i e x_j , assumindo que $w(x_i, x_j) \in [0,1]$ e $\sum_{x_j \in \Omega^3} w(x_i, x_j) = 1$, sendo Ω^3 a grelha tridimensional da imagem [75].

Na definição clássica, o filtro considera que cada voxel pode ser ligado a todos os outros mas, por razões computacionais, o número de voxéis avaliados na média pesada pode ser limitada a um volume de procura V_i centrado em x_i [75].

Para cada voxel x_j em V_i , a distância euclidiana com peso gaussiano $\| \cdot \|_{2,a}^2$, definida por Buades et al., é calculada entre $u(N_i)$ e $u(N_j)$. Esta distância é a L2 clássica, convolvida com um *kernel* gaussiano de desvio padrão a , e mede a distância entre as intensidades das vizinhanças. Dada esta distância, o peso $w(x_i, x_j)$ é dado por:

$$w(x_i, x_j) = \frac{1}{Z_i} e^{-\frac{\|u(N_i) - u(N_j)\|_{2,a}^2}{h^2}} \quad (4.2)$$

onde Z_i é uma constante de normalização para assegurar que $\sum_{x_j \in \Omega^3} w(x_i, x_j) = 1$ e h atua como um parâmetro de suavização, controlando o decaimento da função exponencial. Quando h é muito elevado, todos os voxéis x_j e V_i terão o mesmo peso $w(x_i, x_j)$ relativo a x_i . O valor restaurado $NL(u)(x_i)$ será a média da intensidade dos voxéis em V_i , levando a uma forte suavização da imagem. Por outro lado, quando h é muito baixo, o decaimento da função exponencial será elevado, pelo que apenas alguns voxéis x_j em V_i , com $u(N_j)$ muito semelhante a $u(N_i)$, terão um peso significativo. Assim, o valor restaurado $NL(u)(x_i)$ tenderá para a média de voxéis com valores de vizinhança similares aos de x_i , levando a uma suavização leve da imagem [75].

O filtro não local proposto por Coupé et al. introduz uma forma automática de ajustar o parâmetro de suavização h e sugere uma implementação por blocos, em conjunto

com uma pré-seleção dos vóxeis mais relevantes, de forma a diminuir a carga computacional que o método tradicional requer [75].

A adaptação do filtro de média não local para ruído Rician sugerida por Wiest-Darsse et al. foi conseguida alterando a $NL(u)x_i = \sum_{x_j \in \Omega^3} w(x_i, x_j)u(x_j)$ (4.1) para:

$$NL_R(u)x_i = \sqrt{\left(\sum_{x_j \in \Omega^3} w(x_i, x_j)u(x_j)^2\right) - 2\sigma^2} \quad (4.3)$$

onde σ^2 é a variância do ruído [76].

4.2.2. Correção das não-homogeneidades

O campo de não-homogeneidades manifesta-se como uma variação suave da intensidade de um tecido homogéneo ao longo da imagem, o que altera o valor da sua intensidade, dependendo da sua localização. Esta não uniformidade de intensidades deve-se ao facto do MRI assumir a existência de um campo magnético B_0 completamente contínuo. Na verdade, B_0 está sujeito a várias alterações, normalmente, devido a fraca uniformidade da radiofrequência da bobina, indução de correntes elétricas (correntes de Eddy) devido a gradientes magnéticos e a anatomia do paciente. Exemplos de imagens afetadas por este efeito estão presentes na Figura 4.2 (a) e (d). Embora estas variações se situem entre os 10% e 20%, tendo pouco impacto no diagnóstico visual, a performance de técnicas de segmentação automática podem ser significativamente degradada, uma vez que é assumido uma homogeneidade de intensidades dentro de cada classe. A remoção deste efeito é conseguida através da modelização do seu comportamento na imagem. Geralmente, o mesmo é assumido como multiplicativo [77].

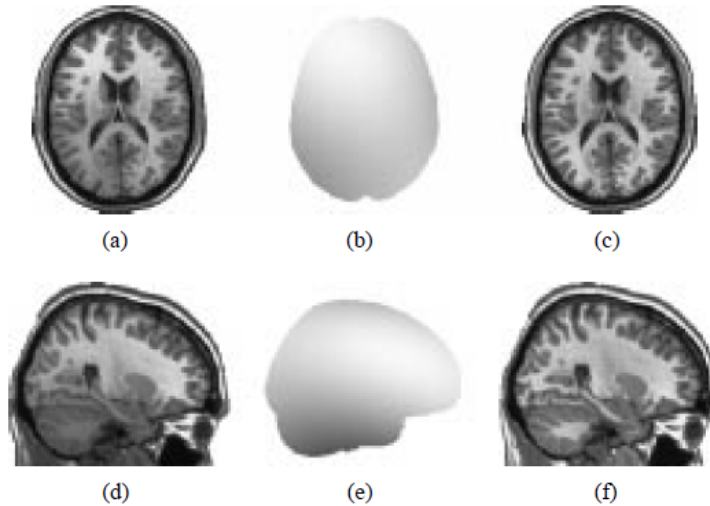


Figura 4.2- Correção da não-homogeneidades de imagens de MR T1-w: (a) e (d) vista axial e sagital de imagens não homogêneas; (b) e (e) campo não uniforme estimado pelo método N3; (c) e (f) imagens corrigidas [77].

O método N3 (*Nonparametric Nonuniform intensity Normalization*) é um dos métodos com melhores resultados na remoção das não-homogeneidades, tendo sido usado em diferentes aplicações [78]–[80]. Este método considera que a imagem é formada pelo seguinte modelo multiplicativo:

$$v(x) = u(x)f(x) + n(x) \quad (4.4)$$

onde na localização x , $v(x)$ é o sinal medido, $u(x)$ é o verdadeiro valor do sinal emitido pelo tecido, $f(x)$ é o campo de variação de intensidades desconhecido e n é o ruído gaussiano, independente de u . O problema de compensar a não-uniformidade da intensidade é a tarefa de estimar f [77].

Considerando o caso livre de ruído, a aplicação do logaritmo à expressão torna o modelo aditivo. Assim, usando a notação $\hat{u}(x) = \log[u(x)]$, o modelo da imagem torna-se [77]:

$$\hat{v}(x) = \hat{u}(x) + \hat{f}(x). \quad (4.5)$$

Sejam U, V e F as funções de densidade de \hat{u} , \hat{v} e \hat{f} , respectivamente. Fazendo a aproximação que \hat{u} e \hat{f} são variáveis aleatórias independentes e não correlacionadas, a distribuição da sua soma é encontrada pela seguinte convolução :

$$V(\hat{v}) = F(\hat{v}) * U(\hat{v}) = \int F(\hat{v} - \hat{f})U(\hat{f}) d\hat{f} \quad (4.6)$$

pelo que a distribuição F pode ser vista como uma suavização da distribuição de intensidades de U . Como \hat{f} é um campo suave de variação lenta, F pode ser estimado iterativamente, de forma a maximizar a frequência de U [77]. Um exemplo da aplicação deste algoritmo está visível na Figura 4.2.

4.2.3. Registo linear para o espaço estereotáxico

Coupé et al. usa o método proposto por Collins et al. para registar linearmente as imagens para o atlas MNI-ICBM152 [28].

O problema de registo consiste na procura da transformação que melhor alinha a imagem de referência I_r e uma outra imagem I_f . Isto é formulado como um problema matemático, através de uma função de custo, $C(I_1, I_2)$, que quantifica a qualidade do registo, e depois encontra a transformação T^* que fornece o custo mínimo $T^* = \operatorname{argmin}_{T \in S_T} C(I_r, I_f \circ T)$ onde S_T é o espaço de transformações permitidas, e $I_f \circ T$ representa a imagem I_f depois de ser transformada por T [81].

No início da formulação do problema, é necessário decidir qual será o espaço permitido de transformações, S_T . Uma classe básica são as transformações lineares que, em 3D, têm 12 graus de liberdade: três parâmetros de translação, três de rotação, três de escala e três de *shear*. Os parâmetros de *shear* causam transformações “trapezoidais” na imagem; por exemplo, um *shear xy* causa uma mudança horizontal na imagem que depende da localização de y . Isto está ilustrado na Figura 4.3, onde um *shear xy* foi aplicado [37], [81].

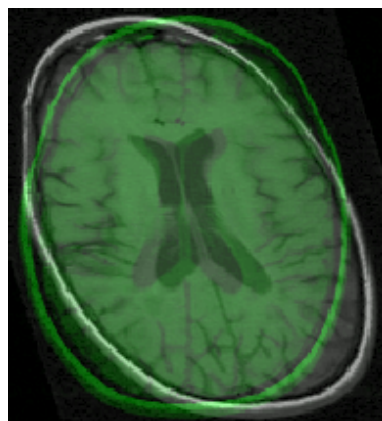


Figura 4.3- Ilustração de uma aplicação de um shear xy a uma imagem RM no plano axial [82].

Usando a formulação por matriz, o ponto $(x, y, z, 1)$ é mapeado para o ponto $(x', y', z', 1)$ usando a transformação:

$$(x', y', z', 1) = [A](x, y, z, 1) \quad (4.7)$$

onde A é uma matriz de transformação 4×4 , contendo doze elementos independentes, e resulta da concatenação de quatro matrizes representates da translação Sh , rotação Sc , escala R e *shear* T [37]:

$$A = [Sh][Sc][R][T]. \quad (4.8)$$

O método de registo criado por Collins et al. utiliza apenas nove parâmetros, não realizando qualquer transformação de *shear* [37].

A maioria dos algoritmos de registo começa pela extração de características das imagens, que serão usadas para avaliar o processo de alinhamento. Devido às transformações efetuadas, as características devem ser insensíveis a pequenas perturbações na forma e robustas na presença de ruído. Para além disso, estas devem ser intrínsecas ao objeto, e não devem depender da posição nem da orientação deste na imagem. Por outras palavras, a característica detetada num dado ponto deve assumir o mesmo valor se o objeto for movido ou rodado na imagem. Collins et al. utiliza duas características: a intensidade da imagem desfocada e magnitude do gradiente. Estas podem ser calculadas em cada vóxel de um volume 3D e verificaram-se suficientes para o registo linear. Uma vez que a desfocagem gaussiana não é independente da escala, as características necessitam de ser recalculadas a cada nova escala do problema de registo [37].

Relativamente ao problema de otimização, Collins et al. utiliza diferentes resoluções espaciais de forma a minimizar problemas associados com mínimos locais da função de custo, pelo que o registo começa com imagens muito desfocadas, sendo mais detalhe adicionado a cada passo [37].

4.2.4. Normalização das intensidades

Um dos problemas para os algoritmos de segmentação em imagens de RM é o facto das intensidades medidas não terem um significado absoluto. Assim, um processo de normalização de intensidades torna-se fundamental para poder assegurar que o contraste de cada tipo de tecido seja consistente para todos os sujeitos da base de dados. Para este

fim, Coupé et al. definiu as intensidades entre o intervalo [0-100] e normalizou todas as imagens com o método proposto por Nyul e Udupa [28].

Em 1999, Nyúl e Udupa propuseram um novo método capaz de transformar o histograma de uma imagem num histograma padrão. Assim, as intensidades dos tecidos passam a ser correspondentes em diferentes imagens, mantendo as relações de intensidade pré-existentes dentro de cada imagem individual [83].

A normalização é conseguida por transformação da moda e da gama de interesse da imagem para valores obtidos a partir uma ou mais imagens de referência. A definição de uma gama de interesse é necessária devido à existência de *outliers*. Assim, o percentil superior e inferior (p_1 e p_2) são definidos, reduzindo a gama de intensidades real $[m_1, m_2]$. Os p_1 e p_2 obtidos a partir das imagens de referência denominam-se s_1 e s_2 . Quando as imagens apresentam uma distribuição bimodal, a moda principal corresponde ao *background*, pelo que a moda μ usada, corresponde à moda secundária do histograma. Em imagens com histogramas unimodais, o parâmetro μ pode ser substituído pelo “final” da única curva, ω , onde o declive da curva é -1. Uma representação destes parâmetros é visível na Figura 4.4 [83].

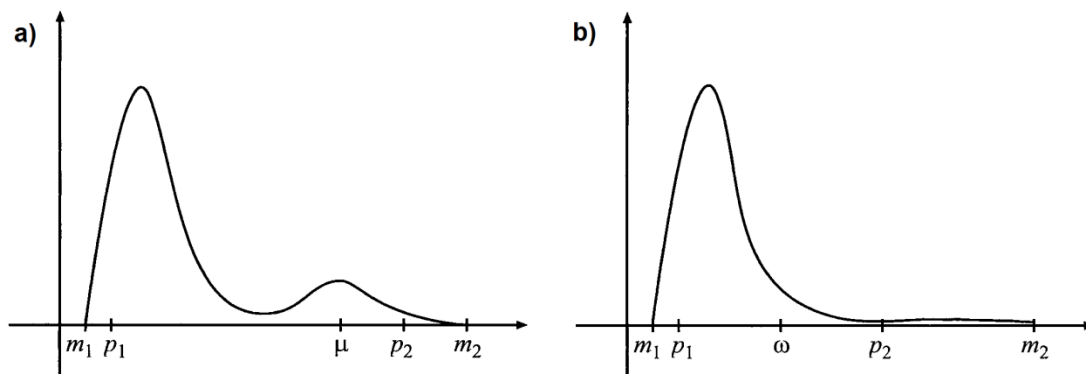


Figura 4.4- Localização dos parâmetros de normalização: a) numa imagem com distribuição bimodal; b) numa imagem com distribuição unimodal [83].

Para a obtenção dos parâmetros, também chamada de fase de treino, as intensidades de cada imagem de referência j , compreendidas entre p_{1j} e p_{2j} , são mapeadas para o intervalo $[s_1, s_2]$, de acordo com [83]:

$$x' = s_1 + \frac{s - p_{1j}}{p_{2j} - p_{1j}}(s_2 - s_1). \quad (4.9)$$

O parâmetro μ_s é o valor médio das modas das imagens de treino, depois de transformadas pela Equação 4.9 [83].

A transformada $\tau(X)$, responsável pela normalização das intensidades de uma nova imagem i , é então conseguida pela equação[83]:

$$\tau(X) = \begin{cases} \mu_s + (x - \mu_i) \frac{\mu_1 - \mu_s}{\mu_{1i} - \mu_i}, & m_{1i} \leq x \leq \mu_i \\ \mu_s + (x - \mu_i) \frac{s_2 - \mu_s}{p_{2i} - \mu_i}, & \mu_i \leq x \leq m_{2i} \end{cases} \quad (4.10)$$

Como é visível pelos intervalos da equação 4.10, os *outliers* da imagem i também são normalizados [83].

4.3. Estratégia de procura na base de dados

A estratégia de procura dentro da biblioteca de imagens destina-se não só, a encontrar *patches* idênticas, mas também a evitar computações desnecessárias. Primeiro, a segmentação é restringida a uma máscara de inicialização. Segundo, é considerado que *patches* similares são encontrados em sujeitos idênticos. Considera-se ainda que a variabilidade entre sujeitos no espaço estereotáxico limitada, pelo que, o volume de procura de *patches* é limitado em torno da localização sob estudo. Finalmente, uma pré-avaliação da similaridade dos *patches* é conseguida pela comparação dos seus contrastes e intensidades [28].

4.3.1. Máscara de inicialização

A segmentação da totalidade da imagem pode ser evitada com definição de uma máscara de inicialização em volta da estrutura de interesse. Um grande número de abordagens pode ser usado como, por exemplo, a dilatação morfológica da máscara do sujeito da base de dados mais próximo da imagem em estudo. Coupé et al. optaram por uma estratégia simples e rápida, utilizando a união de todas as segmentações manuais da base de dados como máscara inicial [28].

4.3.2. Seleção dos sujeitos

Uma seleção ao nível dos sujeitos é efetuada, de uma forma que se aproxima à seleção dos melhores sujeitos no método de fusão de *labels* [72]. Neste método, Coupé et al. usam a soma dos quadrados das diferenças (SQD) ao longo da máscara de inicialização. Esta estratégia foi escolhida porque a SQD é sensível a variações de contraste e intensidade, pelo que, é esperado encontrar um maior número de *patches* idênticas em sujeitos com menores SQDs. O mesmo conjunto dos N elementos mais similares é mantido durante toda a segmentação [28].

4.3.3. Pré-seleção de *patches*

Finalmente, tal como proposto para questões de filtragem [76], Coupé et al. realizam uma pré-seleção dos *patches*, que serão comparados de forma a reduzir o tempo de computação. Usando medidas estatísticas simples, como a média e a variância, é possível descartar, *a priori*, a maioria dos *patches* pouco relevantes. Neste método, são usados o contraste e a intensidade como critério de pré-seleção. Baseado no primeiro e no segundo termo de uma medida de similaridade estrutural (SSIM), a medida de pré-seleção pode ser escrita como:

$$SS = \frac{2\mu_i\mu_{s,j}}{\mu_i^2 + \mu_{s,j}^2} \times \frac{2\sigma_i\sigma_{s,j}}{\sigma_i^2 + \sigma_{s,j}^2} \quad (4.11)$$

onde μ representa a média e σ o desvio padrão dos *patches* centrados no vóxel x_i (vóxel em consideração) e no vóxel $x_{s,j}$ (vóxel na localização j do sujeito s). Se o valor de ss for maior que um dado *threshold*, a distância de intensidades entre os *patches* i e j é calculada [28].

Finalmente, o método de procura proposto limita os *patches* considerados a um volume de procura, isto é, apenas *patches* próximas do vóxel em estudo, em termos de espaço estereotáxico, são consideradas. Para além da redução drástica do tempo de computação, a pré-seleção efetuada também reduz o número de *outliers* na fusão não local dos *labels* [28].

4.4. Fusão de *labels* por média não local

A estratégia de fusão de *labels* proposta por Coupé et al. é baseada no estimador de média não local proposto por Buades et al., em 2005 [28], [70]. Neste método, a distância baseada na intensidade entre os *patches* é usada para calcular uma média pesada das amostras [28].

4.4.1. Estimador de média não local

Para todos os vóxeis x_i da imagem a ser segmentada, a estimativa do *label* final é baseada na fusão de *labels* pesada $v(x_i)$ de todas as amostras classificadas dentro do volume de procura V_i para o N sujeitos selecionados:

$$v(x_i) = \frac{\sum_{s=1}^N \sum_{j \in V_i} w(x_i, x_{s,j}) y_{s,j}}{\sum_{s=1}^N \sum_{j \in V_i} w(x_i, x_{s,j})} \quad (4.12)$$

onde $y_{s,j}$ é o *label* atribuído por um especialista ao vóxel $x_{s,j}$ na localização j do sujeito s e $w(x_i, x_{s,j})$ é o peso atribuído a $y_{s,j}$ por comparação de *patches*. Dependendo da similaridade entre o *patch* envolvendo x_i e o *patch* envolvendo $x_{s,j}$, o peso é calculado como:

$$w(x_i, x_{s,j}) = \begin{cases} \exp\left(\frac{-\|P(x_i) - P(x_{s,j})\|_2^2}{h}\right) & ss \geq th \\ 0, & ss \leq th \end{cases} \quad (4.13)$$

onde $P(x_i)$ representa o *patch* cúbico centrado em x_i , $\|\cdot\|_2$ é a distância L2 normalizada (pelo número de elementos) calculada entre cada intensidade dos elementos dos *patches* $P(x_i)$ e $P(x_{s,j})$ e h é um parâmetro de decaimento. Como explicado na secção 4.3.3, se a similaridade estrutural ss entre os *patches* é menor que o *threshold* th , então o peso não é calculado, assumindo automaticamente o valor 0 [28].

Finalmente, considerando os *labels* y definidos em $\{0,1\}$, a classificação final $L(x_i)$ é dada por:

$$L(x_i) = \begin{cases} 1, & v(x_i) \geq 0,5 \\ 0, & v(x_i) < 0,5 \end{cases} \quad (4.14)$$

O método também pode ser aplicado com *labels* probabilísticos definidos em $[0,1]$, sem quaisquer modificações [28]. A Figura 4.5 exemplifica, com uma pré-seleção de três sujeitos, os passos envolvidos na segmentação do vóxel x_i [28].

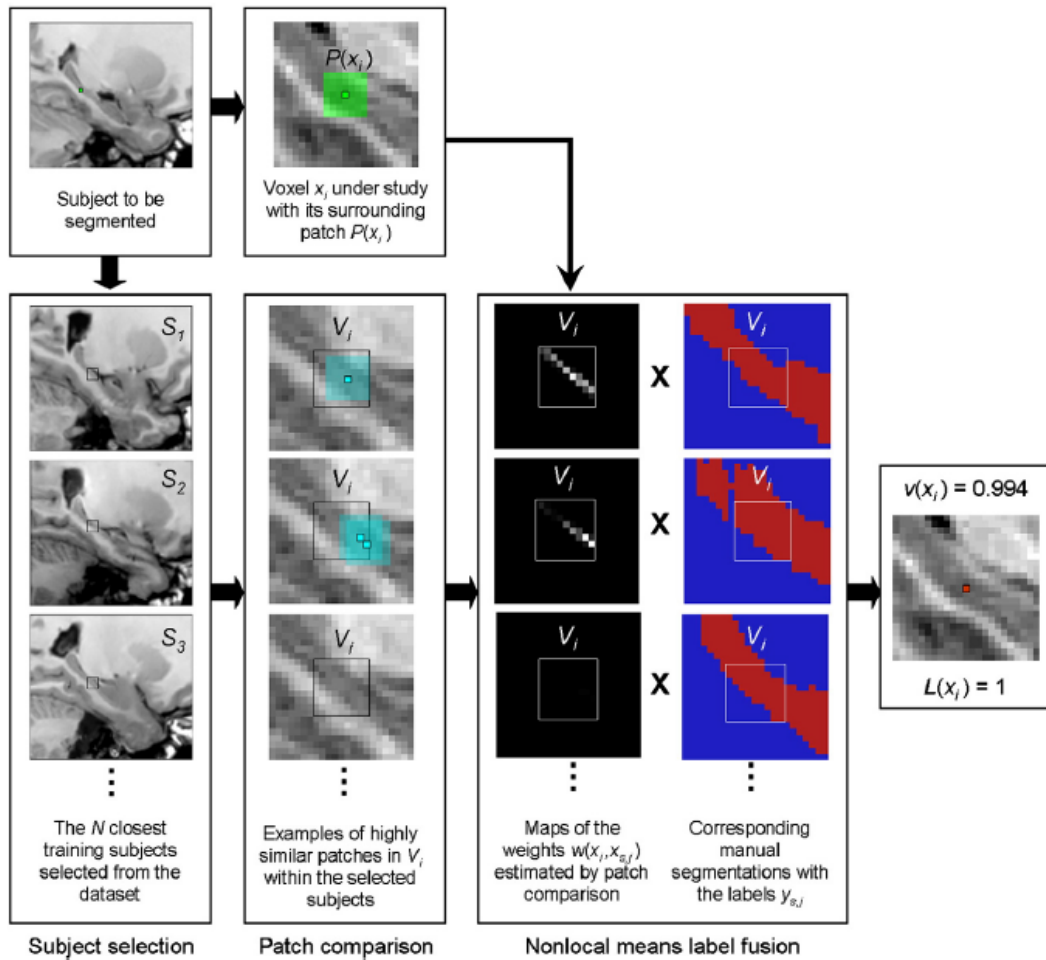


Figura 4.5- Visão geral dos diferentes passos envolvidos na segmentação do vóxel x_i incluído na máscara de inicialização. O *patch* $P(x_i)$ é comparado com todos os *patches* $P(x_{s,j})$ contidos no volume de procura V_i no interior dos N sujeitos selecionados (neste caso, $N = 3$). O mapa de pesos mostra que os maiores pesos são atribuídos aos *patches* $P(x_{s,j})$ mais similares a $P(x_i)$. Depois da fusão de média não local dos *labels* $y_{s,j}$, a estimativa resultante é $v(x_i) = 0.994$, pelo que o label final é $L(x_i) = 1$ [28].

4.4.2. Adaptação local de h

Em problemas de estimação, o ajuste do parâmetro de decaimento h desempenha um papel crucial. Como já foi referido no caso da filtragem não local, se h for muito baixo, apenas algumas amostras serão tomadas em consideração, quando h é muito elevado, todas as amostras tendem a ter o mesmo peso e a estimação aproxima-se da média clássica [28].

Idealmente, o valor h deve depender da distância de intensidades do *patch* sob consideração e o conteúdo da base de dados. De facto, quando a biblioteca contém *patches* muito parecidas com o *patch* em estudo, h necessita de ser diminuído de forma a reduzir a influência de outros *patches*. Contudo, quando não existem *patches* similares, h precisa de ser aumentado de forma a relaxar a seleção. Para alcançar esta adaptação local de h automaticamente, Coupé et al. sugeriu uma estimacão $h(x_i)$ baseada na distância mínima entre $P(x_i)$ e $P(x_{j,s})$:

$$h(x_i) = \lambda^2 \times \arg_{x_{j,s}} \min \|P(x_i) - P(x_{j,s})\|_2^2 + \varepsilon, \quad (4.15)$$

onde ε é uma pequena constante para assegurar estabilidade numérica no caso do *patch* sob consideração estar contido na biblioteca, e λ é uma constante. Este tipo de adaptação local tem sido similarmente usado no denoising adaptativo de MRI [28].

4.5. Sumário

O algoritmo proposto por Coupé et al. segue uma abordagem semelhante aos métodos baseados em atlas, usando imagens de RM segmentadas manualmente por especialistas como antecedentes, de forma a segmentar estruturas anatómicas. No entanto, o método usa uma escala menor através de *patches* e a fusão de *labels* é feita ponderadamente para cada vóxel, com uma abordagem não local.

A variabilidade causada pela formação da imagem é minimizada através de um extenso passo de pré-processamento: filtragem, correção das não-homogeneidades, transformação linear para o espaço estereotáxico e normalização das intensidades.

Uma vez que o processamento não local é computacionalmente dispendioso, algumas medidas são necessárias, de forma a evitar computações desnecessárias e a acelerar a execução do método. Primeiro, a segmentação é restringida a uma máscara de inicialização. É ainda considerado que *patches* similares são encontrados em sujeitos idênticos, havendo uma pré-seleção de sujeitos. Para além disso, uma vez que a variabilidade entre sujeitos no espaço estereotáxico limitado, o volume de procura de *patches* é limitado em torno da localização sob estudo. Finalmente, uma pré-avaliação da similaridade dos *patches* é conseguida pela comparação dos seus contrastes e intensidades.

Capítulo 5 Segmentação e gradação simultânea de estruturas anatómicas

Em 2012, Coupé et al. propôs um novo método capaz da extração de biomarcadores de Alzheimer a partir de imagens de ressonância magnética. O algoritmo SNIPE (*Scoring by Nonlocal Patch Estimator*) é uma extensão da segmentação baseada em *patches* apresentada no Capítulo 4 **Error! Reference source not found.** [66].

Através do uso de populações de treino com estados clínicos diferentes, isto é, sujeitos de controlo (CN) e pacientes com doença de Alzheimer (AD), o estimador de média não local é utilizado para avaliar a proximidade de cada vóxel da imagem RM em estudo com as populações de treino. Uma vez que a estimativa da gradação e a fusão de *labels* exige o mesmo passo de comparação de *patches*, segmentação e gradação das estruturas estudadas podem ser alcançadas simultaneamente, sem necessidade de computação adicional. O valor médio da gradação obtida foi proposto como um novo biomarcador para prever o estado clínico dos pacientes em estudo [66].

Do ponto de vista clínico, a capacidade de prever Alzheimer, isto é, distinguir pacientes com MCI progressivo ou estático, é mais importante do que detetar AD. Contudo, a previsão é claramente mais desafiante uma vez que, para além das mudanças anatómicas serem mais subtis, o conjunto de sujeitos com MCI é muito heterogéneo, sendo que alguns indivíduos convertem para Alzheimer (pMCI) e outros não (sMCI). A distinção destas duas classes é crucial para a proposta de um novo biomarcador. Assim, o método foi usado para segmentar e classificar o hipocampo (HC) e o córtex entorrinal (CE), tendo a sua capacidade de separação de estados clínicos sido testada em imagens de RM da base de dados da *Alzheimer's Disease Neuroimaging Initiative* (ADNI) [60], [66].

Neste capítulo é descrita a evolução do método de segmentação de Coupé et al. (apresentado no Capítulo 4) para o algoritmo SNIPE. É efetuada uma descrição do conjunto imagens usado, bem como das alterações nos passos de pré-processamento. Para além disso, os biomarcadores finais propostos por Coupé et al. são apresentados em conjunto com o esquema de validação usado para comparar as capacidades de

discriminação entre estados clínicos. Por fim, os resultados obtidos pelo método para todas as imagens da base de dados são também exibidos.

5.1. Da segmentação para a gradação

Como já foi referido, o algoritmo SNIPE é uma extensão do método de segmentação apresentado no Capítulo 4 . Este propõe uma nova forma de estimar a aproximação de uma imagem com duas populações com estados clínicos diferentes (AD e CN). Para isto, Coupé et al. introduz um novo conceito de “grau”, baseado em *patches*, que reflete a aproximação do *patch* em torno do vóxel em estudo com todos os *patches* das populações de treino. Desta forma, a informação da vizinhança é usada para guiar, robustamente, a procura dos padrões que são específicos de um dado subconjunto da base de dados. Quando as populações de treino incluem informação de sujeitos com diferentes estados clínicos, desta abordagem resulta, para cada vóxel, uma estimativa da gradação, isto é, um grau de aproximação para um grupo ou para outro [66].

A gradação, $g(x_i)$, é obtida com uma adaptação da Equação 4.12:

$$g(x_i) = \frac{\sum_{s=1}^N \sum_{j \in V_i} w(x_i, x_{s,j}) \cdot p_s}{\sum_{s=1}^N \sum_{j \in V_i} w(x_i, x_{s,j})} \quad (5.1)$$

onde p_s é um label correspondente estado clínico do sujeito de treino s . No caso de Coupé et al. $p_s = -1$ foi usado para pacientes com AD e $p_s = 1$ para sujeitos CN. Assim, um valor da gradação $g(x_i)$ negativo indica que a vizinhança do vóxel x_i é mais característica de AD, enquanto que, uma gradação $g(x_i)$ positiva aponta para uma vizinhança de x_i mais próxima de CN. O valor absoluto $\|g(x_i)\|$ providencia a confiança dada na estimação da gradação. Quando $\|g(x_i)\|$ é próximo de zero, o método indica que o *patch* em estudo tem similaridades com as duas populações, fornecendo pouca informação discriminatória. Quando $\|g(x_i)\|$ é próximo de 1, o método deteta uma proximidade elevada entre o *patch* em estudo e os *patches* presentes numa das populações. Finalmente, para cada sujeito, um valor da gradação médio é calculado sobre todos os vóxeis da segmentação da estrutura, ou seja, todos os vóxeis onde $v(x_i) > 0.5$. Uma vez que a segmentação e gradação envolvem o mesmo passo de comparação de *patches*, a gradação pode ser efetuada ao mesmo tempo

que a segmentação é estimada [66]. Uma ilustração do funcionamento do algoritmo SNIPE está presente na Figura 5.1.

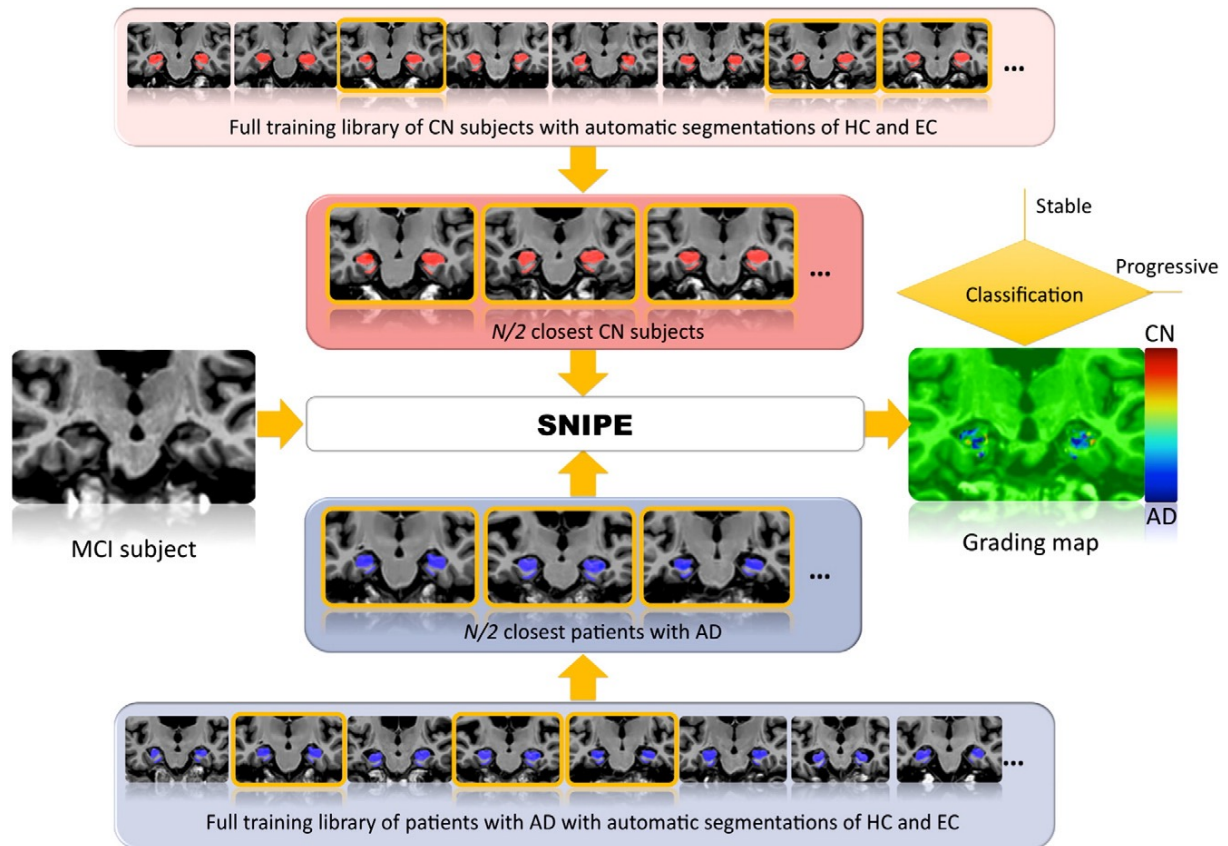


Figura 5.1- Exemplo da execução do SNIPE para um sujeito com MCI. A biblioteca de treino é usada para estimar uma segmentação e um mapa da gradação, que podem ser usados como biomarcadores para o diagnóstico de Alzheimer [60].

Várias estratégias podem ser usadas para obter o valor médio da gradação das estruturas estudadas. Primeiro, cada lado da estrutura pode ser usado separadamente. Segundo, é possível atribuir o mesmo peso ao HC e CE esquerdo e direito, como por exemplo, $\overline{g}_{HC} = \frac{(\overline{g}_{HC_{esquerdo}} + \overline{g}_{HC_{direito}})}{2}$. Esta estratégia, de fundir os dois lados das estruturas para aumentar a robustez a imprecisões de segmentação, foi proposta por Chupin et al. num estudo volumétrico [66], [84]. Embora Coupé et al. tenha verificado resultados semelhantes para análises independentes do HC e CE, para o complexo HC-CE, a melhor estratégia foi encontrada com a média da gradação esquerda e direita, ao longo de todo o complexo (dando maior importância ao hipocampo devido ao seu maior volume), e depois calcular a média dos dois lados ($\overline{g}_{HC-CE} = \frac{(\overline{g}_{HC-CE_{esquerdo}} + \overline{g}_{HC-CE_{direito}})}{2}$) [66].

5.2. Base de dados ADNI: 834 imagens de base

O estudo de Coupé et al. teve como objetivo investigar a capacidade do algoritmo SNIPE diagnosticar precocemente AD. Para isso, 834 imagens de base T1w com 1,5 T da base de dados ADNI foram usadas. Os scans foram divididos em quatro populações, sendo os sujeitos com MCI considerados pMCI se converterem para AD até julho de 2011. Desta construção de população resultou um conjunto de dados com 231 CN, 238 sMCI, 167 pMCI e 198 AD. Os quatro grupos são os propostos por Wolz et al. [60], [85].

5.3. Pré-processamento

Tal como no método de segmentação, a variabilidade causada pela formação da imagem é minimizada com um passo de pré-processamento. Da mesma forma, todas as imagens foram filtradas e sofreram uma correção das não homogeneidades com os métodos descritos em 4.2.1. e em 4.2.2. , respetivamente [60].

O registo das imagens para o mesmo sistema de coordenadas continua a ser um passo crucial do algoritmo. No SNIPE, o registo é igualmente realizado linearmente, com o método proposto por Collins et al.. No entanto, em vez do uso do atlas MNI-ICBM152, um *template* específico para a população é derivado a partir da base de dados ADNI, com o algoritmo publicado por Fonov et al. em 2011 [60].

Os restantes passos do pré-processamento mantêm-se, mas com uma exceção. A normalização das intensidades com método Nyul e Udupa é precedida pela extração do cérebro com o algoritmo BEaST (*Brain extraction based on nonlocal segmentation technique*), proposto por Eskildsen et al. em 2012 [60].

5.3.1. Criação de um atlas específico para a população

Com os novos métodos de MRI dos últimos 20 anos, vários grupos de investigação reportaram mudanças relacionadas com a idade nos volumes de substância cinzenta e branca de estruturas subcorticais e na extensão da mielinização. Logo, uma variabilidade significativa tem sido encontrada em dados volumétricos e metabólicos ao longo de populações e géneros [86].

Várias técnicas automáticas de registo, classificação de tecidos, e análise estatística usam *templates* de cérebros, como o mni_autoreg de Collins et al., o SPM de Ashburner e Friston, e o FSL de Smith et al. [37], [87], [88]. Contudo, estas técnicas não são ideais para estudos em populações específicas, já que estes atlas são obtidos por cálculo da média de MRIs de adultos saudáveis. Para estudos pediátricos, por exemplo, o cérebro em desenvolvimento não é uma simples versão diminuída de um cérebro adulto, pelo que o uso destes *templates* pode introduzir erros na análise. Em 2000, por exemplo, Muzik et al. mostrou que, usando um *template* de adultos com o SPM, o registo pediátrico é mais variável que o registo de MRI de adultos [86].

O método de criação de *templates* específicos descrito por Fonov et al. depende do mecanismo de registo não linear do algoritmo ANIMAL. Para estimar a transformação não linear requerida entre as imagens de origem e o volume alvo (o atlas ICBM152 original), o ANIMAL tenta corresponder, hierarquicamente, características das imagens através da maximização da correlação das intensidades entre as imagens de origem e as imagens alvo. Primeiro, as deformações necessárias para corresponder de versões desfocadas dos dados de origem e alvo são estimadas, produzindo um campo de deformações 3D. De seguida, o campo de deformações é interpolado e usado como entrada na próxima iteração do procedimento, onde a desfocagem é diminuída e a estimação do campo de deformação é aperfeiçoada. Desta forma, deformações grandes e suaves são recuperadas primeiro, e deformações mais pequenas e locais são recuperadas em último [86].

Essencialmente, o problema pode ser formulado do seguinte modo: dado um conjunto de n volumes MRI em 3D (I_1, \dots, I_n) , o objetivo do método é encontrar um *template* 3D Φ , que satisfaz duas restrições simultaneamente, uma de intensidade, e outra de transformação. A primeira restrição pretende minimizar a diferença média das intensidades quadradas entre o *template* Φ e cada sujeito I_i , transformado para corresponder ao *template*:

$$\Phi^* = \arg \min_{\Phi} \left[\sum_{i=1}^n \int (\Phi(v) - I_i(\Psi_{i,\Phi}(v)))^2 dv \right], \quad (5.2)$$

onde v é uma coordenada de volume, $\Psi_{i,\Phi}$ são os mapas 3D individuais formados a partir do *template* Φ para cada sujeito I_i , $\Phi(v)$ é a intensidade do *template* na localização v , e $I_i(\Psi_{i,\Phi}(v))$ é a intensidade no MRI do sujeito i depois de transformada por $\Psi_{i,\Phi}$. A

transformação Ψ é restringida através um de simples modelo elástico, de tal forma que para cada sujeito i [86]:

$$\Psi_{i,\phi} = \arg \min_{\phi} \int_{\text{volume}} (\Phi(v) - I_i(\Psi(v)))^2 dv. \quad (5.3)$$

A segunda restrição serve para minimizar a magnitude de todas as deformações $\Psi_{i,\phi}$ necessárias para transformar o *template* Φ para cada sujeito i :

$$\Phi^* = \arg \min_{\phi} \left[\sum_{i=1}^n \int |\Psi_{i,\phi}(v) - v|^2 dv \right] \quad (5.4)$$

Resumindo, as Equações 59 e 60 são minimizadas simultaneamente; a Equação 60 é minimizada para cada par sujeito-*template* [86].

A transformação $\Psi_{i,T}$ é representada por um campo vetorial de deformação h , de forma a que $\Psi(x) = x + h(x)$ e $\Psi^{-1}(x) = x + h^{-1}(x)$, onde $h(x)$ pode ser definido numa grelha discreta com uma dada distância (tamanho do *step*) entre nós, tal como o algoritmo ANIMAL. Seguindo este formalismo, Fonov et al. propôs um novo algoritmo iterativo para minimizar as duas funções, primeiro a diferença da média ao quadrado em termos de deformações (Equação 60) e segundo, a diferença da média ao quadrado em termos de intensidades (Equações 59). A denotação $X_{i,k}$ é atribuída ao mapeamento de cada sujeito encontrado a cada passo consecutivo do algoritmo; T_k é a aproximação atual; quando o algoritmo converge $X_{i,k} \rightarrow \Psi_{i,\phi}, T_k \rightarrow \Phi^*$, Φ^* e $\Psi_{i,\phi}$ representam, respetivamente, o *template* de mínima deformação e a transformação final. Assim, o algoritmo pode ser descrito como se segue [86]:

- 1) Dado T_k (a aproximação do *template* Φ^* na iteração k), para cada scan I_i , o $X_{i,k}$ é calculado (mapeamento do *template* para cada scan individual i , na iteração k) usando $Y^{-1}_{i,k-1}$ (mapeamentos inversos corrigidos do scan i , na iteração $k-1$) como ponto de partida. A transformada identidade foi usada para a primeira iteração e o ICBM152 linear usado como T_0 .
- 2) Cálculo do erro residual baseado na deformação média $X_{0,k}$ do *template* atual T_k :

$$X_{0,k} = x + \sum_n h_i(x) / n \quad (5.5)$$

- 3) Cálculo do mapeamento inverso corrigido: $Y_{i,k} = X^{-1}_{i,k} \circ X_{0,k}$, onde \circ indica a composição das transformações. Este passo corresponde à função de minimização da

Equação 60. É de realçar que $Y_{i,k}$ é definida no espaço de cada sujeito, e deve ser invertida numericamente para poder ser usada.

- 4) Aplicação dos mapeamentos inversos corrigidos a cada sujeito e gerar uma média que será usada como novo *template*, deste modo minimizando a Equações 59:

$$T_{k+1}(x) = \sum_n I(Y_{i,k}(x)) / n \quad (5.6)$$

- 5) Repetir o passo 1 até uma convergência ser alcançada.

No registo não linear, o processo é repetido com uma diminuição do tamanho dos *steps* de uma forma hierárquica. Para a condição de convergência, o valor quadrático médio da deformação residual média do campo vetorial gerado no segundo passo é calculado, e o processo é parado assim que a diferença entre dois passos consecutivos seja menor que um dado limite. Uma representação esquemática do algoritmo está presente na Figura 5.2 [86].

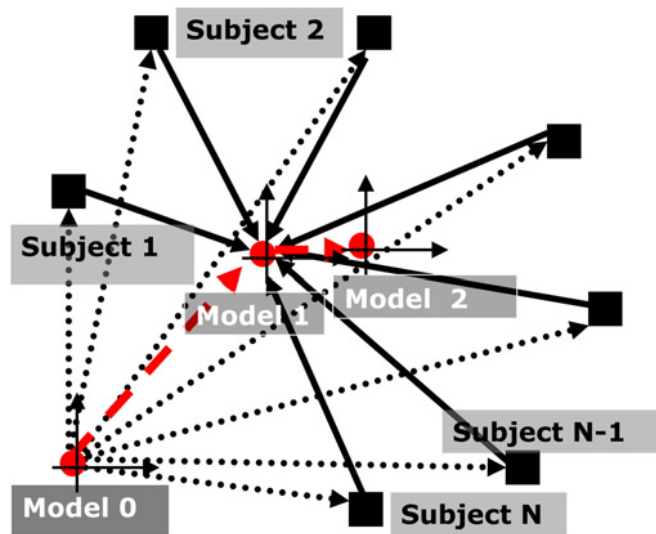


Figura 5.2- Representação esquemática do algoritmo de construção do modelo: as linhas pontilhadas representam o mapeamento de um voxel no modelo inicial (modelo 0) para cada sujeito; linhas sólidas representam o mapeamento de sujeitos individuais para o modelo seguinte (modelo 1); linhas tracejadas representam o erro residual de cada voxel do modelo a cada iteração [86].

5.3.2. Extração do cérebro

A extração do cérebro é um passo importante em várias análises neuroimagingológicas, como a classificação de tecido e a segmentação. Por exemplo, a falha na remoção da dura-

máter (a maior das meninges) pode levar a uma medida exagerada da espessura cortical, enquanto a remoção de parte do cérebro levaria a uma subestimação [89].

Um método de extração preciso deve excluir todos os tecidos externos ao cérebro, como o crânio, dura-máter, olhos sem remover qualquer parte do cérebro. No entanto não existe uma única norma definida para o que deve ou não ser excluído. O algoritmo BEaST opta pela remoção de todos os tecidos extra cerebrais que se aproximam das intensidades da SB e SC. Seguindo este princípio, o LCR próximo do cérebro não é excluído, uma vez que nas imagens T1w este pode ser facilmente diferenciado das estruturas não líquidas [89].

O método BEaST é inspirado na segmentação baseada em *patches* proposta por Coupé et al.. Da mesma forma, o SQD é usado como medida para a estimação de similaridades entre *patches*. O uso do SQD requer que a intensidade dos tecidos cerebrais seja consistente ao longo dos sujeitos e sequências de imagem. Assim, a correção das não-homogeneidades e a normalização das intensidades são passos essenciais na construção da biblioteca [89].

O princípio de segmentação usado neste algoritmo já foi descrito na presente dissertação, no entanto, de forma a otimizar a performance da extração cerebral, o tamanho dos *patches* usados necessitam de ser maiores do que os usados na segmentação de estruturas menores. Por exemplo, um *patch* reduzido da zona da dura-máter pode ser facilmente confundido por SC em imagens T1w. Por conseguinte, um *patch* de dimensões elevadas, incluindo mais informação estrutural, é necessário para evitar inclusão de tecido extracerebral. Isto não é computacionalmente prático na resolução estereotáxica. Assim, uma *framework* com várias resoluções é usada, mantendo o método computacionalmente exequível sem abdicar do uso de *patches* com grandes dimensões [89]. Resumidamente, um esquema de várias resoluções permite a propagação da segmentação ao longo de diferentes escalas, através do uso de segmentações resultantes da escala anterior como inicialização da segmentação na escala atual [89].

A base de dados de imagens, máscara de inicialização e a imagem alvo na resolução estereotáxica V^j são reamostradas a uma resolução menor V^{j-k} , e a segmentação é realizada. O estimador de média não local $v_{V^{j-k}}(x_i)$, com a resolução V^{j-k} , é propagado para uma maior resolução V^{j-k+1} usando interpolação trilinear. Um valor de confiança α

para atribuir *labels* aos vóxeis estudados é definido para cada escala. Vóxeis com $v_{V^{j-k}}(x_i) < \alpha$ ou $v_{V^{j-k}}(x_i) > 1 - \alpha$ são classificados como background ou objeto, respetivamente. A segmentação destes dois conjuntos de vóxeis é considerada final e não sofrem processamento adicional. Vóxeis cujo $v_{V^{j-k}}(x_i)$ se situa no intervalo $[\alpha; 1 - \alpha]$ são propagados para uma resolução mais elevada, V^{j-k+1} . Este procedimento é repetido até a resolução estereotáxica V^j ser alcançada, onde o threshold do estimador passa a ser 0,5 [89].

5.4. SNIPE - *Scoring by Nonlocal Patch Estimator*

Como já foi referido, a gradação das estruturas em estudo pelo algoritmo SNIPE são conseguidas por estimativa não local da similaridade do sujeito em consideração com diferentes populações de treino. Com o esquema não local, o SNIPE é capaz de lidar com a variabilidade entre os sujeitos ao explorar a redundância de informação no corpo humano. Para além disso, com princípio de comparação por *patches*, o método é capaz de detetar mudanças anatómicas subtis causadas pela doença [60].

Depois do pré-processamento, o diagnóstico do paciente com o SNIPE pode ser dividido em três passos: propagação dos labels, segmentação e gradação da estrutura e a classificação final [60].

5.4.1. Propagação dos *labels*

Quando se lida com um elevado número de imagens, a obtenção da segmentação manual para a totalidade da base de dados torna-se impraticável. Logo, o primeiro passo do algoritmo SNIPE é a propagação de um número limitado de segmentações manuais para a totalidade da biblioteca de treino. No estudo de Coupé et al., apenas as populações de AD e CN foram usadas como base para a classificação; por isso, a propagação dos *labels* é apenas efetuada para os sujeitos CN e AD. Assim, 20 imagens aleatórias foram selecionadas das populações AD e CN (10 CN e 10 AD) e segmentadas manualmente no espaço estereotáxico, por especialistas, usando o protocolo descrito por Pruessner et al. [90]. De seguida, as segmentações manuais foram usadas para segmentar a totalidade das populações de AD e CN. Como resultado, segmentações anatómicas estiveram disponíveis para os 231 sujeitos

CN e 198 pacientes com AD, sendo usadas para a construção da biblioteca de treino [60], [66].

5.4.2. Segmentação e gradação

Logo que o passo de propagação de *labels* termine, a biblioteca de treino resultante pode ser usado pelo SNIPE para estimar mapas de segmentação e gradação de qualquer imagem [60], [66]. Exemplos dos mapas da gradação estão apresentados na *Figura 5.3*.

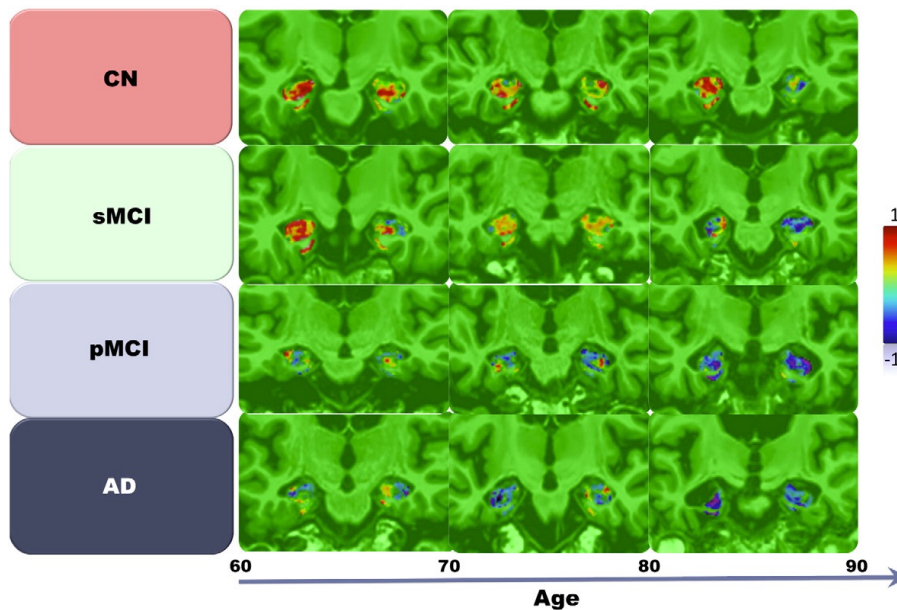


Figura 5.3- Exemplos de mapas da gradação típicos para cada população de acordo com a idade [60].

O primeiro passo da segmentação é a seleção dos sujeitos. Os $N/2$ sujeitos mais próximos de cada população de treino (isto é, AD e CN) são conseguidos pelo uso do SQD ao longo da máscara de inicialização [60], [66].

Com os sujeitos já escolhidos, o patch envolvente de cada vóxel (incluído na máscara de inicialização) da imagem em estudo é comparado com todos os *patches* dos N *templates* de treino, selecionados das populações de AD e CN. Assim, uma segmentação e um mapa da gradação são obtidos simultaneamente [60], [66].

O passo final é o uso das estimativas obtidas para calcular o volume total e o valor médio da gradação da estrutura. Estes serão os biomarcadores usados na classificação do paciente [60], [66].

5.4.3. Gradação

O passo da gradação é baseado na análise discriminante linear. Em 2012, Coupé et al. mostrou que, uma exatidão ligeiramente mais elevada podia ser alcançada na classificação AD vs. CN, usando uma análise discriminante quadrática; contudo, para permitir a comparação com resultados até à data publicados, a análise discriminante linear (LDA) foi usada no estudo. Para além disso, melhores precisões foram obtidas com a introdução da idade do paciente, em conjunto com o volume ou a gradação; assim, todos os resultados apresentados por Coupé et al. foram obtidos usando as idades dos sujeitos como característica adicional na análise [60].

5.5. Esquema de validação

O esquema de validação é usado para comparar a capacidade dos biomarcadores baseados no SNIPE de discriminar as diferentes populações. Os biomarcadores estudados foram o volume do HC, a gradação do HC, o volume do EC, a gradação do EC e as suas combinações [60], [66].

Primeiro, a segmentação e a gradação das estruturas é efetuada para todas as imagens da base de dados. A seleção dos *templates* é feita depois da remoção do sujeito em estudo da biblioteca de treino. Para um dado sujeito, os N *templates* de treino mais próximos são selecionados dos restantes sujeitos da base de dados. Depois, a segmentação e a gradação foram obtidas, assegurando que nenhum sujeito é incluído no seu próprio processamento. A ausência dos sujeitos MCI é implícita, uma vez que apenas as populações de AD e CN são usadas como *templates* de treino [60].

Uma vez obtidos os volumes e classificações para todos os sujeitos, uma análise discriminante linear com uma abordagem LOOCV (*leave-one-out cross-validation*) foi realizada. Neste tipo de validação, o classificador é treinado com $n - 1$ amostras e depois usado para classificar a amostra restante. A especificidade, sensibilidade e exatidão são apresentadas para cada teste [60], [66].

5.6. Resultados

A Tabela 5.1 apresenta as medidas de sucesso da separação das diferentes populações obtida pelos biomarcadores imagiológicos propostos por Coupé et al..

De uma forma geral, estes resultados mostram que os biomarcadores baseados na gradação superam os biomarcadores baseados no volume. Para além disso, os biomarcadores baseados no EC são menos eficientes do que os biomarcadores baseados no HC. Finalmente, a combinação do volume e da gradação não melhora significativamente os resultados obtidos pela gradação isoladamente [60].

Tabela 5.1- Resultados obtidos com os diferentes biomarcadores para a diferenciação das diferentes populações. Estes foram obtidos usando a análise discriminante linear com um procedimento LOOCV. Os valores na tabela correspondem à exatidão (Exa),sensibilidade (Sen) e a especificidade (Esp) em percentagem (%) [60]

	Exa%	Sen%	Esp%	Exa %	Sen%	Esp%	Exa %	Sen%	Esp%
AD vs. CN	HC			EC			HC-EC		
Volume	79	76	82	70	68	72	78	76	80
Gradação	88	83	92	83	75	90	89	84	93
Volume+ Gradação	87	83	91	83	74	91	88	81	92
pMCI vs. CN	HC			EC			HC-EC		
Volume	75	73	76	69	66	71	75	74	75
Gradação	85	80	88	79	73	83	86	80	89
Volume+ Gradação	85	80	88	80	73	85	85	80	88
AD vs. sMCI	HC			EC			HC-EC		
Volume	68	67	70	62	57	66	69	67	70
Gradação	73	71	75	72	69	74	77	77	78
Volume+ Gradação	73	71	75	73	70	75	77	77	77
pMCI vs. sMCI	HC			EC			HC-EC		
Volume	62	61	63	59	59	59	63	63	64
Gradação	71	70	71	66	62	68	70	69	71
Volume+ Gradação	71	70	72	65	60	68	70	71	69
AD vs. pMCI	HC			EC			HC-EC		
Volume	56	51	59	51	48	54	55	51	58
Gradação	58	57	60	62	63	60	60	60	59
Volume+ Gradação	58	57	59	61	63	59	60	61	59
sMCI vs. CN	HC			EC			HC-EC		
Volume	63	65	62	60	65	55	64	65	63
Gradação	69	74	63	63	68	58	68	76	60

Volume+ Gradação	69	76	62	64	72	55	69	76	63
------------------	----	----	----	----	----	----	----	----	----

Análises individuais de cada distinção de populações bem como estudos sobre o impacto da idade nos biomarcadores são efetuados em [60] mas, por questões de brevidade, não são apresentados nesta dissertação.

5.7. Sumário

O algoritmo SNIPE é uma extensão da segmentação baseada em *patches* apresentada no Capítulo 4. Através do uso de populações de treino com estados clínicos diferentes, o estimador de média não local é utilizado para avaliar a proximidade de cada vóxel da imagem RM em estudo com as populações de treino. A informação da vizinhança é usada para guiar, robustamente, a procura dos padrões que são específicos de um dado subconjunto da base de dados. Quando as populações de treino incluem informação de sujeitos com diferentes estados clínicos, desta abordagem resulta, para cada vóxel, um grau de aproximação para um grupo ou para outro.

Tal como no método de segmentação, a variabilidade causada pela formação da imagem é minimizada com um passo de pré-processamento. Assim, todas as imagens sofreram uma filtragem, correção das não-homogeneidades, transformação linear para o espaço estereotáxico usando um *template* específico derivado a partir da base de dados ADNI, extração do cérebro e normalização das intensidades.

Depois do pré-processamento, o diagnóstico do paciente pode ser dividido em três passos: propagação dos labels, segmentação e gradação da estrutura e a classificação final.

O esquema de validação é usado para comparar a capacidade dos biomarcadores, extraídos do hipocampo e do córtex entorrinal, de discriminar as diferentes populações. A segmentação e a gradação das estruturas são efetuadas para todas as imagens da base de dados, assegurando que nenhum sujeito é incluído no seu próprio processamento. Uma vez obtidos os volumes e classificações para todos os sujeitos, as classificações são obtidas através de uma análise discriminante linear com uma abordagem LOOCV.

Capítulo 6 Trabalho Experimental

Como já foi referido em 3.3. , o método SNIPE foi escolhido como a base para a criação de uma ferramenta de CAD de Alzheimer. Como nenhuma versão do algoritmo SNIPE se encontra disponível publicamente, uma das principais tarefas da dissertação foi a implementação da totalidade do método. Além disso, uma nova forma de calcular o peso usado na fusão dos *labels* é proposta. O SNIPE baseado no filtro Laplaciano (SNIPE-LoG) usa os mapas das bordas, calculados pelo filtro de Laplaciano, como uma medida adicional, na comparação entre *patches*.

A Figura 6.1 apresenta um esquemático dos dois métodos implementados. Diagramas mais detalhados das implementações do SNIPE e do SNIPE-LoG estão apresentados, na Figura 6.4 e na Figura 6.6, respetivamente.

Esta secção descreve os problemas e soluções encontrados na implementação do algoritmo. Enuncia ainda, os algoritmos de pré-processamento usados, justifica a biblioteca de processamento de imagem escolhida e descreve a base de dados usada. Uma vez que uma versão do algoritmo SNIPE foi implementada, as principais escolhas, como a linguagem de programação e estruturas de dados usadas, foram fundamentadas. Por fim, uma alteração ao algoritmo SNIPE é ainda proposta, onde informação das bordas das imagens é considerada no processamento.

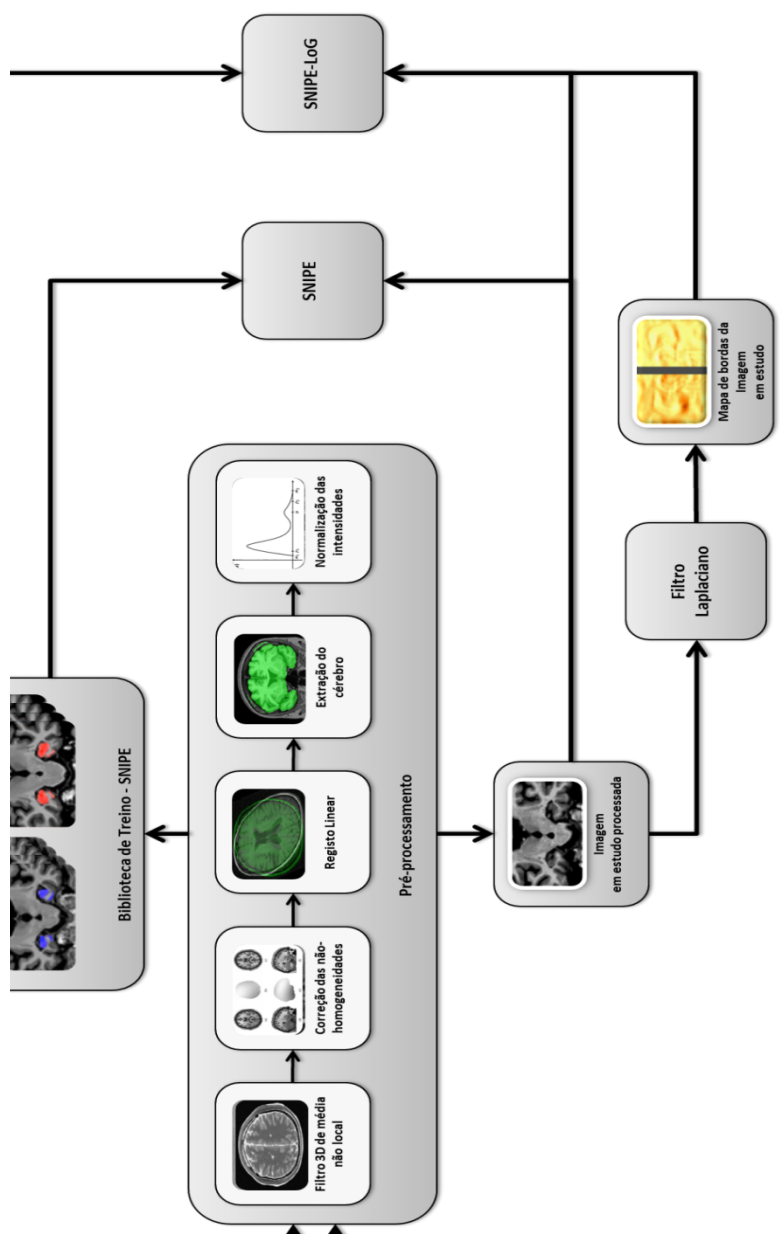


Figura 6.1- Esquema do trabalho experimental efetuado. Os blocos a branco representam métodos usados disponibilizados por bibliotecas externas. Os blocos cinzentos representam métodos ou processos implementados durante a dissertação.

6.1. Pré-processamento

Tal como descrito no 5.3. , a implementação do algoritmo SNIPE requer um extenso passo de pré-processamento de forma a minimizar a variabilidade entre as imagens da base de dados.

A maioria dos passos de pré-processamento necessários, como o registo linear e a normalização das intensidades, têm implementações inseridas em vários *softwares* disponíveis à comunidade científica, como é o caso do Freesurfer e do ITK (*Insight Toolkit*) [91], [92]. No entanto, o primeiro passo do pré-processamento, a filtragem com o filtro de média não local, encontra-se disponível publicamente apenas no pacote de ferramentas MINC (*Medical Image NETCDF*). Assim, este foi o *software* escolhido para aplicar o pré-processamento a todas as imagens da base de dados.

A ferramenta *mincnlm* é uma adaptação do código original para o MINC realizada por Pierrick Coupé e Pierre Yger. Todos os parâmetros padrão são os sugeridos em [75] pelo que nenhuma alteração dos parâmetros foi efetuada[91]. O passo seguinte, a correção das não-homogeneidades, foi efetuado com a ferramenta *nu_correct*. Esta implementa o algoritmo N3, apresentado anteriormente em 4.2.2. [91].

Uma grande quantidade dos processamentos seguintes pode ser efetuado com recurso a um dos *pipelines* disponibilizados no pacote MINC, o *standard_pipeline*. Este *script* realiza, em série, o registo da imagem para o espaço de um modelo fornecido, a normalização linear das intensidades para a intensidade do *template* e a extração do cérebro com o algoritmo BEeST [91]. Tal como referido em 5.3.1. , o trabalho mais recente de Coupé et al. regista as imagens para um modelo de população específica derivado da base de dados ADNI. No entanto, a construção de uma modelo a partir de um número tão elevado de imagens requer um *cluster* com grande poder computacional, algo que não esteve disponível para esta dissertação. Assim, tal como numa das primeiras publicações de Coupé et al. [66], o atlas ICBM152 foi usado como modelo. A transformação de cada imagem é aplicada à segmentação manual respetiva, mantendo as máscaras no mesmo espaço que a imagem correspondente. Um exemplo de uma imagem e da máscara respetiva transformadas para o espaço estereotáxico está visível na Figura 6.22.

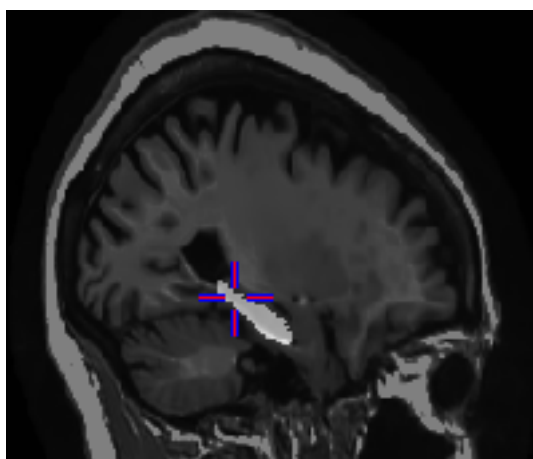


Figura 6.2- Slice sagital de uma imagem e a respetiva máscara hipocampal no espaço estereotáxico.

Por fim, a normalização das intensidades entre os sujeitos foi executada com a ferramenta `minc_nuyl`, uma implementação do algoritmo publicado por Nyul e Udupa [83]. Nesta execução, a fase de treino é realizada com uma única imagem fornecida como parâmetro. Como Coupé et al. não especifica o modelo usado na sua normalização, a imagem com maior contraste foi escolhida. Esta foi encontrada pela diferença do 2º e 98º percentil, de modo a excluir possíveis *outliers* na consideração.

6.2. Base de dados ADNI1

O ADNI é um estudo longitudinal, ainda em desenvolvimento e realizado em colaboração com vários centros de investigação, para desenvolver biomarcadores bioquímicos, genéticos, imagiológicos e clínicos para a deteção prematura e o acompanhamento de Alzheimer [93].

Descobertas científicas e evoluções tecnológicas ao longo do tempo levaram a ajustes nos protocolos utilizados. Assim, embora todos os estudos tenham o objetivo de avançar a investigação do Alzheimer, três diferentes fases do projeto foram definidas: ADNI1, ADNIGO e ADNI2. A primeira fase, o ADNI1, começou em 2004 e incluiu 400 sujeitos diagnosticados com MCI, 200 sujeitos com AD e 200 idosos de controlo. Em 2009, ADNI1 foi estendida com o ADNIGO o qual avaliou o grupo do ADNI1 e adicionou 200 participantes em estados iniciais de MCI. Em 2011, o ADNI2 incorporou 550 novos participantes no estudo [93]. A Figura 6.33 ilustra a duração de cada uma das fases [94].

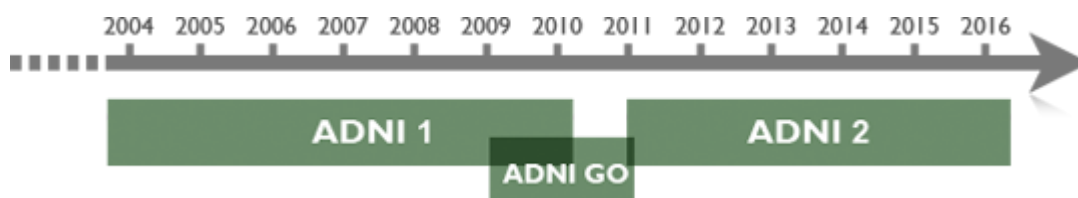


Figura 6.3- Cronograma ilustrativo das durações das diferentes fases do ADNI.

Com um elevado número de investigadores a trabalhar com o mesmo conjunto de dados, existe a possibilidade de comparações diretas entre algoritmos de análise ou passos de pré-processamento de imagem. No caso do ADNI1, apenas 819 dos sujeitos que tiveram a avaliação de base (testes neurofisiológicos, imagiológicos e às amostras de fluído retiradas) cumpriram os critérios de inclusão do manual de procedimentos do ADNI e foram considerados como a amostra oficial do ADNI1 [93]. Critérios de inclusão/exclusão detalhados estão descritos em [95]. Outros agrupamentos de dados foram também sugeridos onde, por exemplo, para promover um maior rigor na análise e comparações significativas entre os diferentes algoritmos de processamento de imagem, o núcleo de MRI do ADNI criou conjuntos de análise padronizados constituídos por imagens que reúnem requerimentos mínimos de controlo de qualidade [96].

Os resultados publicados pelo algoritmo SNIPE (exibidos em 5.6.) foram obtidos com a utilização de todas as 834 T1w MRIs de base disponíveis até à data. Para reproduzir estes resultados, as mesmas imagens deveriam ser usadas. Porém, tendo em conta como a informação das imagens está apresentada na base de dados, isto não foi conseguido. Embora a data de aquisição de cada imagem seja fornecida, as primeiras 834 imagens ordenadas por data de exame não correspondem ao grupo descrito por Coupé et al. (231 CN, 238 sMCI, 167 pMCI e 198 AD). Assim, as 819 imagens de base dos pacientes do ADNI1 foram usadas nesta dissertação para a construção da biblioteca de treino. Isto permite aproximar o conjunto de dados com os usados pelo SNIPE, e comparações diretas com algoritmos que usem o mesmo grupo padronizado.

Todas as 819 T1w MRIs dos sujeitos alistados passaram pelo *pipeline* de processamento descrito em 6.1. . Embora nenhum especialista tenha sido consultado para inspecionar a qualidade das imagens resultantes, a imagem de um dos sujeitos MCI foi removida do estudo subsequente, devido a erros na conversão de formatos que levaram a

uma falha total no processamento dessa imagem. O resultado do processamento deste sujeito é apresentado no Anexo I – Falha do pré-processamento.

Mais uma vez, de forma a minimizar variações nos resultados apresentados por Coupé et al., o mesmo critério foi usado para a construção das diferentes quatro populações. Os diagnósticos são efetuados com base em resultados MMSE, CDR (*Clinical Dementia Rating*) e cumprem os critérios NINCDS-ADRDA para Alzheimer provável. Assim, sujeitos com MCI foram considerados progressivos se existir uma conversão para Alzheimer até Julho de 2011 [85]. A Tabela 6.1 apresenta os detalhes demográficos das populações resultantes.

Tabela 6.1- Detalhes demográficos dos sujeitos usados divididos por população com diagnósticos até 2011. Incluí o número de sujeitos pertencentes (Nº de sujeitos), a percentagem de população masculina (% Masculina), a idade mais o desvio padrão (Idade $\pm \sigma$) e o resultado do MMSE mais o desvio padrão (MMSE $\pm \sigma$)

População	Nº de sujeitos	% Masculina	Idade $\pm \sigma$	MMSE $\pm \sigma$
CN	229	52,7	76,0 \pm 7,0	29,1 \pm 1,0
sMCI	231	52,5	74,8 \pm 7,7	26,3 \pm 3,3
pMCI	165	0,51	74,5 \pm 7,1	26,7 \pm 3,6
AD	193	52,8	75,4 \pm 7,4	23,4 \pm 2,1

Uma vez que um grande número dos sujeitos com MCI do ADNI foi integrado nas fases seguintes do estudo, desde de 2011, nova informação útil foi divulgada. A Tabela 6.2 apresenta as distribuições dos sujeitos pelas populações, considerando os diagnósticos efetuados ao longo das diferentes fases do ADNI.

Tabela 6.2- Detalhes demográficos dos sujeitos usados divididos por população com diagnósticos até a atualidade. Incluí o número de sujeitos pertencentes (Nº de sujeitos), a percentagem de população masculina (% Masculina), a idade mais o desvio padrão (Idade $\pm \sigma$) e o resultado do MMSE mais o desvio padrão (MMSE $\pm \sigma$)

População	Nº de sujeitos	% Masculina	Idade $\pm \sigma$	MMSE $\pm \sigma$
CN	229	52,7	76,0 \pm 7,0	29,1 \pm 1,0
sMCI	201	51,7	75,2 \pm 7,6	26,3 \pm 3,3
pMCI	195	51,2	74,5 \pm 7,1	26,7 \pm 3,6
AD	193	52,8	75,4 \pm 7,4	23,4 \pm 2,1

A Tabela 6.2 mostra que, desde de 2011, um grande número dos pacientes com MCI converteram para AD.

6.3. SNIPE - *Scoring by Nonlocal Patch Estimator*

Nesta secção serão explicadas as escolhas relativas à linguagem de programação, bibliotecas utilizadas e passos da implementação do algoritmo.

6.3.1. Linguagem de programação

Uma das primeiras decisões tomadas foi a linguagem de programação a ser usada. A implementação de Coupé et al. foi programada em C-MEX para o Matlab, onde partes do código são escritas em C e compiladas antes do Matlab correr o restante código [28], [97]. Embora esta abordagem aumente significativamente a performance comparativamente com implementações em Matlab “puro”, linguagens de programação mais poderosas podem ser utilizadas para contornar um dos principais problemas do método: o tempo de computação. O C++ é uma destas linguagens.

O principal objetivo do C++ é providenciar as características da programação por orientação a objetos, numa linguagem em C, sem compromisso de eficiência. A estrutura orientada a objetos permite o reuso de código, o que pode reduzir o tempo de implementação, e facilita a sua manutenção ou alteração. No entanto, geralmente falando, a principal vantagem do C++ é a sua performance, quando comparada com outras linguagens de alto nível [98]. Por esta razão, o C++ foi a linguagem escolhida para a implementação do algoritmo. Esta linguagem é umas das linguagens mais utilizadas na atualidade, apesar de ser considerada uma linguagem de difícil aprendizagem [97].

6.3.2. Biblioteca de processamento de imagem

Qualquer processamento de imagem requer a leitura da informação presente na imagem, neste caso, os valores de intensidade dos vóxeis das imagens de RM. Para isto, várias bibliotecas encontram-se disponíveis para C++. O ITK, por exemplo, é uma das ferramentas mais usadas para segmentação e registo de imagens médicas e tem a capacidade de ler alguns dos formatos de MRI mais comuns, como o DICOM e Analyze [92].

No entanto, o pré-processamento descrito em 6.1. foi realizado com as ferramentas do MINC, exigindo o uso de imagens nesse mesmo formato. Para evitar conversões de formato adicionais a biblioteca de C++ do MINC foi escolhida para leitura das imagens. Apesar da documentação disponível ser substancialmente inferior, a biblioteca do MINC tem uma instalação mais fácil e rápida do que a biblioteca do ITK.

Com a biblioteca MINC, o primeiro passo da leitura das imagens é o `miopen_volume`, uma função capaz abrir um volume MINC existente a partir de uma diretoria fornecida e associá-lo à variável introduzida como parâmetro. Esta variável representa o objeto MINC e permite a leitura das intensidades dos vóxeis das imagens [91].

Na primeira implementação do algoritmo, as intensidades dos vóxeis foram lidas com a função `miget_real_value`. Esta função extrai os valores de uma posição num volume MINC [91]. A localização é definida por um *array* fornecido como parâmetro. Embora esta função tenha permitido o correto funcionamento do algoritmo, o seu elevado tempo de execução tornou a sua utilização impraticável.

Numa nova versão da implementação, os valores de intensidade foram lidos com a função `miget_real_value_hyperslab`. Esta função permite retirar uma porção contígua de dados do volume e armazená-los num *array* unidimensional. O *hyperslab* é um conceito central da biblioteca MINC e permite processamentos muito mais rápidos do que a obtenção individual de cada vóxel, embora requeira uma maior quantidade de memória [91]. Nos casos onde a vizinhança é considerada no processamento, como na abordagem por patches, a conversão dos dados do volume 3D para um *array* unidimensional obriga o utilizador a criar funções extra que traduzam vizinhanças 3D para o espaço unidimensional.

6.3.3. Propagação dos *labels*

Como referido em 5.4.1. , a segmentação manual de um grande número de imagens de RM não é uma tarefa prática. Este é um trabalho altamente especializado que pode chegar a ocupar 2 horas para segmentação de apenas uma estrutura [39]. Coupé et al. contornou este problema com a utilização do algoritmo SNIPE para propagar os *labels* de um número limitado de segmentações manuais para as restantes imagens da base de dados [60]. Nesta fase, a biblioteca de treino foi composta por 20 imagens aleatórias seleccionadas

das populações AD e CN (10 CN e 10 AD). As segmentações do hipocampo e do córtex entorrinal foram obtidas manualmente com o protocolo descrito em [90]. Uma vez que o acesso a profissionais especializados capazes de segmentar estruturas cerebrais é limitado, a mesma abordagem não foi seguida nesta dissertação.

Desde da publicação do algoritmo SNIPE, o ADNI lançou um novo projeto, envolvendo os maiores especialistas internacionais em segmentação hipocampal em pesquisa de Alzheimer. O principal objetivo foi a harmonização dos protocolos existentes para a segmentação manual do hipocampo. Um consenso foi conseguido resultando num protocolo onde o hipocampo é considerado com a inclusão do álveos/fimbria, de toda a cauda hipocampal e a fronteira medial do corpo de acordo com a morfologia visível. Esta definição captura 100% do tecido do hipocampo, 100% da atrofia relacionada com o Alzheimer e demonstra boa confiança em estimativas preliminares intra (0,98) e inter (0,94) técnico [99].

Com o protocolo definido em [99], o ADNI disponibilizou, para as suas imagens, centenas de segmentações manuais do hipocampo. Para o conjunto de imagem base do ADNI1 especificadas em 6.2. , 239 máscaras hipocampais estão disponibilizadas, das quais 102 são de sujeitos com AD e 137 pertencem à população CN. Estas imagens foram usadas como biblioteca de treino para a segmentação do hipocampo dos restantes sujeitos das populações de AD e CN. Por fim, as segmentações manuais e automáticas de 229 sujeitos CN e 193 sujeitos AD podem ser usadas para a construção da biblioteca de treino usado no algoritmo. Infelizmente, segmentações manuais do córtex entorrinal não são fornecidas pelo ADNI pelo que esta estrutura não será analisada nesta dissertação.

6.3.4. Detalhes de implementação

Nesta dissertação, todos os parâmetros ótimos encontrados em [60] foram usados. Para o hipocampo, as dimensões dos *patches* foram mantidas a $7 \times 7 \times 7$ e a inicialização do volume de procura foi definida com as dimensões $9 \times 9 \times 9$. Contudo, no caso de nenhum *patch* idêntico for encontrado, o raio de procura é aumentado por um vóxel até pelo menos um *patch* similar seja encontrado [66].

O *threshold* th definido foi de 0,95 para todos os testes. Este valor foi escolhido empiricamente pois providencia um bom equilíbrio entre exatidão de segmentação e redução de tempo computacional [28].

A adaptação local automática do parâmetro de suavização $h^2(x_i)$ obteve uma ligeira melhoria nos resultados de segmentação usando $\lambda = 0,5$ em vez de $\lambda = 1$.

Finalmente, os N sujeitos foram igualmente selecionados das duas populações ($N/2$ da população AD e $N/2$ da população de CN) de forma a assegurar que o número de *patches* disponíveis das duas populações seja grosseiramente o mesmo [66]. Por razões computacionais, apenas cerca de 25% da biblioteca ($N_{AD}=50$ e $N_{CN}=50$) foi usada [60].

6.3.5. Implementação

Depois da construção da biblioteca de treino, da escolha da linguagem de programação e da biblioteca de processamento de imagem, o algoritmo SNIPE pôde, então, ser implementado. Como na implementação de qualquer algoritmo, várias direções podem ser tomadas. Aqui serão explicadas as principais abordagens e estruturas de dados escolhidas. A Figura 6.44 expõe o esquema geral da implementação.

O primeiro passo da implementação passa pela criação da máscara de inicialização. Tal como Coupé et al. em [60], utilizou-se uma estratégia simples e rápida para a construção de uma máscara inicial: a união de todas as segmentações manuais da base de dados. Assim, todas as coordenadas do espaço estereotáxico, onde pelo menos uma das segmentações identifica tecido do hipocampo, são guardadas num vetor de inteiros. Uma vez que esta máscara é constante para um dado conjunto de treino, todas as coordenadas são guardadas num ficheiro específico. Isto permite que a máscara inicial seja rapidamente lida a cada utilização do *software*, evitando a necessidade de abrir e percorrer centenas de máscaras hipocampais.

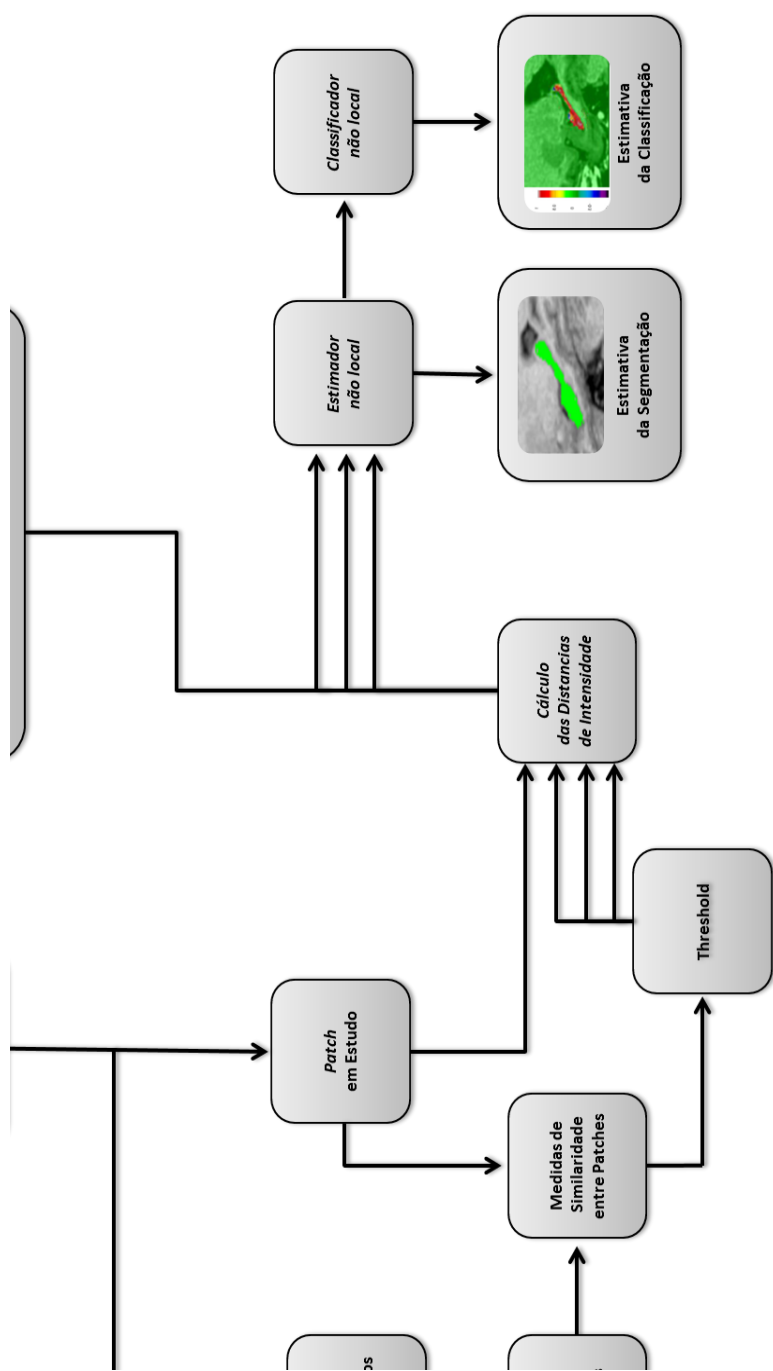


Figura 6.4 - Diagrama representativo da implementação do algoritmo SNIPE.

Para a realização dos testes de validação, a máscara do sujeito em estudo não é utilizada para a criação da máscara inicial. A sobreposição máscara inicial com um sujeito exemplo está presente na Figura 6.55.

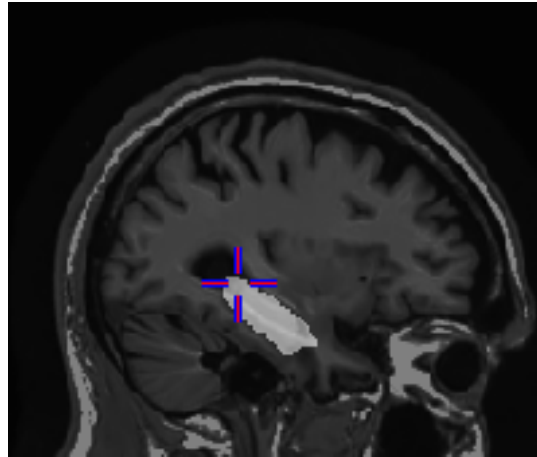


Figura 6.5- Slice sagital de uma imagem sobreposta com união de todas as máscaras da base de dados.

As SQDs entre a imagem em estudo e cada sujeito da base de dados ao longo das coordenadas da máscara de inicialização são então calculadas. Os $N/2$ sujeitos de cada uma das populações com menor SQD são selecionados, sendo os únicos a fornecer informação para o processo de segmentação e gradação que se segue. Como as SQDs dependem da imagem em estudo, estas têm de ser recalculadas cada vez que o algoritmo é executado. As imagens e as máscaras dos N sujeitos são guardadas num vetor de hyperslabs (*arrays*) durante a execução do método, evitando o consumo de tempo necessário para abrir e fechar as imagens de cada vez que uma comparação é efetuada. Como a biblioteca MINC permite abrir porções da imagem, apenas o volume envolvente da máscara inicial é armazenado, reduzindo a memória necessária para a execução do algoritmo.

Tendo em conta a elevada carga computacional exigida pelas abordagens de *patches* não locais, para além de uma seleção dos sujeitos, Coupé et al. recomenda uma pré-seleção dos *patches* considerados dentro do volume de procura [66]. Da mesma forma, a medida de similaridade da $ss = \frac{2\mu_i\mu_{s,j}}{\mu_i^2 + \mu_{s,j}^2} \times \frac{2\sigma_i\sigma_{s,j}}{\sigma_i^2 + \sigma_{s,j}^2}$ (4.11) foi usada como parâmetro de triagem. Visto que as imagens são lidas para um *array* unidimensional, funções capazes de traduzir vizinhanças 3D para o espaço unidimensional tiveram de ser implementadas.

Uma vez que grande parte dos vóxeis da máscara de inicialização se encontram próximos, existe uma grande sobreposição entre os volumes de procura dos diferentes vóxeis. Logo, se as vizinhanças forem consideradas independentes, o método calculará a média e variância dos mesmos *patches* múltiplas vezes, aumentando, desnecessariamente, o tempo de computação. Assim, de forma a evitar cálculo redundante e aumentar a velocidade do algoritmo, os mapas de médias e variâncias locais de todos os sujeitos de treino selecionados são calculados previamente.

Cada vóxel da imagem de estudo presente na máscara de inicialização é analisado individualmente. O *patch* envolvente é comparado com os *patches* dentro do volume de procura provenientes de cada um dos sujeitos selecionados. Se a medida de similaridade for superior a um certo *threshold*, a norma L2 da diferença de intensidades entre *patches* é calculada e armazenada num vetor de *doubles*, em conjunto com o *label* anatómico de treino e o *label* do estado clínico correspondentes. O valor mínimo da distância L2 é guardado para a adaptação local do parâmetro de decaimento h .

Depois das diferenças de intensidades, para todos os *patches* similares, serem calculadas, a $w_{xi, x_{s,j}}$ =
$$\begin{cases} \exp \frac{-\|P(x_i) - P(x_{j,s})\|_2^2}{h} & ss \geq th \\ 0, & ss \leq th \end{cases} \quad (4.13 \text{ pode ser usada para}$$

calcular os pesos finais. Estes, em conjunto com os *labels* anatómicos correspondentes, permitem estimar, com a $v(x_i) = \frac{\sum_{s=1}^N \sum_{j \in V_i} w(x_i, x_{s,j}) y_{s,j}}{\sum_{s=1}^N \sum_{j \in V_i} w(x_i, x_{s,j})} \quad (4.12, \text{ um } label \text{ anatómico}$

para o vóxel em estudo. Se o vóxel for avaliado como pertencente ao hipocampo, uma estimativa da gradação é calculada. Esta usa os mesmos pesos anteriormente calculados pelo que não incrementa, significativamente, o tempo de computação. Por fim, as segmentações e a classificações resultantes são traduzidas para o espaço tridimensional e guardados em duas novas imagens. Estas podem ser usadas para extrair os biomarcadores propostos por Coupé et al em [60].

6.4. SNIPE baseado no filtro Laplaciano

Como já foi referido, em imagens médicas, vários padrões repetidos existem em localizações diferentes da imagem [28]. Estas redundâncias foram exploradas pela primeira

vez, como forma de filtragem, por Buades et al. com o filtro de média não local [70]. Embora o filtro tenha providenciado resultados promissores, uma suavização excessiva foi observada em algumas regiões, resultando na perda de algumas bordas e pequenas estruturas na imagem [100]. Algo similar acontece com o método de segmentação com *patches* não locais proposto por Coupé et al., onde as bordas das estruturas obtêm valores de segmentação menos discriminativos [28].

Em 2013, Bhujle et al. propôs um novo filtro para remoção de ruído Rician em imagens de RM com o objetivo de conservar um maior detalhe em imagens médicas. Neste método, a filtragem de média não local é executada, mas os pesos foram modificados para combinar funções de similaridade de intensidades e de bordas. O filtro Laplaciano (LoG) combina filtragem gaussiana com o operador de Laplace tendo sido usado para a extração das bordas das imagens. O filtro foi avaliado qualitativamente e quantitativamente em imagens simuladas e reais, demonstrando bons resultados de filtragem com uma maior preservação do detalhe [100].

6.4.1. Modificação do cálculo do peso no filtro de média não local

A maioria dos métodos de processamento de MRI baseados em média não local usa o mesmo tipo de estratégia explicada em 4.4. para calcular a distância entre dois *patches* [100]. No entanto, no trabalho de Bhujle et al. em [100], o filtro LoG é aplicado às imagens de forma a extrair as bordas. O filtro LoG é menos sensível à presença de ruído gaussiano devido à filtragem passa-abaixo gaussiana prévia ao cálculo das derivadas. O *kernel* LoG com um parâmetro de espalhamento σ_w tem a forma [100]:

$$LoG(x, y) = -\frac{1}{\pi\sigma_w^4} \left[1 - \frac{x^2+y^2}{2\sigma_w^2} \right] e^{-\frac{(x^2+y^2)}{2\sigma_w^2}}. \quad (6.1)$$

Em [100], $\sigma_w = 1$ foi usado empiricamente e a procura de *patches* similares foi feita ao longo de uma janela 11×11 . As bordas da imagem em estudo foram extraídas apenas uma vez no início do método. A similaridade entre os mapas de bordas do *pixel* em estudo i e o *pixel* a ser comparado j é calculada como [100]:

$$D_1 = \|LoG(x_i) - LoG(x_j)\|_{2,b}. \quad (6.2)$$

$LoG(x_i)$ e $LoG(x_j)$ são as bordas para os *patches* $P(x_i)$ e $P(x_j)$ dos pixels x_i e x_j , respectivamente. O valor de D_1 é calculado como a soma as diferenças absolutas dos elementos dos dois *patches*. Para providenciar um maior peso ao *pixel* central dos *patches*, a função é pesada por um *kernel* Gaussiano com um desvio padrão b . A distância D_1 é posteriormente combinada com a distância calculada pela similaridade de intensidades para formar um novo peso. A distância baseada nas intensidades é a mesma que no método original de média não local [100]:

$$D_2 = \|I(x_i) - I(x_j)\|_{2,b}^2 \quad (6.3)$$

O novo peso é formado depois da combinação das duas distâncias:

$$w_m(i, j) = \frac{\exp\left(-\frac{D_2 + D_1}{h^2}\right)}{\sum_j \exp\left(-\frac{D_2 + D_1}{h^2}\right)} \quad (6.4)$$

onde h é um parâmetro de suavização [100].

6.4.2. Adaptação do peso do SNIPE

A dificuldade do filtro de média não local de conservar pequenos detalhes da imagem é, de certa forma, propagada para o algoritmo SNIPE, uma vez que este tem algumas dificuldades em segmentar regiões com alta variabilidade entre os sujeitos, como o caso das fronteiras das estruturas[28].

Uma vez que o algoritmo SNIPE também é inspirado no filtro de média não local, uma adaptação do filtro de Bhujle et al. pode ser facilmente enunciada. Desta forma, *patches* com bordas similares assumem um maior peso na decisão do algoritmo SNIPE. Assim, usando o mesmo conceito, para o SNIPE a distância D_1 pode ser calculada como:

$$D_1(x_i, x_{j,s}) = \|LoG(x_i) - LoG(x_{j,s})\|_{2,b}, \quad (6.5)$$

onde $LoG(x_i)$ e $LoG(x_{j,s})$ são os mapas das bordas dos *patches* envolventes do vóxel em estudo x_i e do vóxel de x_j do sujeito de treino s . A distância de intensidades é calculada da mesma forma pelo que:

$$D_2(x_i, x_{j,s}) = \|P(x_i) - P(x_{j,s})\|_2^2. \quad (6.6)$$

Finalmente, a introdução da informação das bordas no algoritmo SNIPE é conseguida $x_i, x_s, j = \exp - P x_i - P x_j, s \geq 2h$ & $ss \geq th$ 0, & $ss \leq th$ (4.13) para:

$$w(x_i, x_s, j) = \begin{cases} \exp \frac{-\|D_1(x_i, x_j, s) + D_1(x_i, x_j, s)\|_2^2}{h} & ss \geq th. \\ 0, & ss \leq th \end{cases} \quad (6.7)$$

6.4.3. Detalhes de implementação

Todos os parâmetros relacionados com o método de *patches* não locais foram mantidos iguais aos sugeridos por Coupé et al. em [60]. Estes incluem um tamanho dos *patches* de $7 \times 7 \times 7$, um volume de procura inicial de $9 \times 9 \times 9$, $\lambda = 0,5$ para o cálculo do parâmetro de suavização e o uso de 100 *templates* da biblioteca [60].

Embora Bhujle et al. sugira a sua própria estratégia de pré-seleção de *patches*, onde são incorporados os valores dos mapas de bordas, os *thresholds* usados não são divulgados pelo que, nesta dissertação, esta estratégia não foi usada. Assim, o *threshold* para a seleção dos *patches* para o valor de similaridade manteve-se $th = 0,95$ [66], [100].

Por fim, tal como Bhujle et al., o desvio padrão do *kernel* do filtro Laplaciano foi mantido, com $\sigma_w = 1$ [100].

6.4.4. Implementação

A modificação do peso do algoritmo SNIPE não

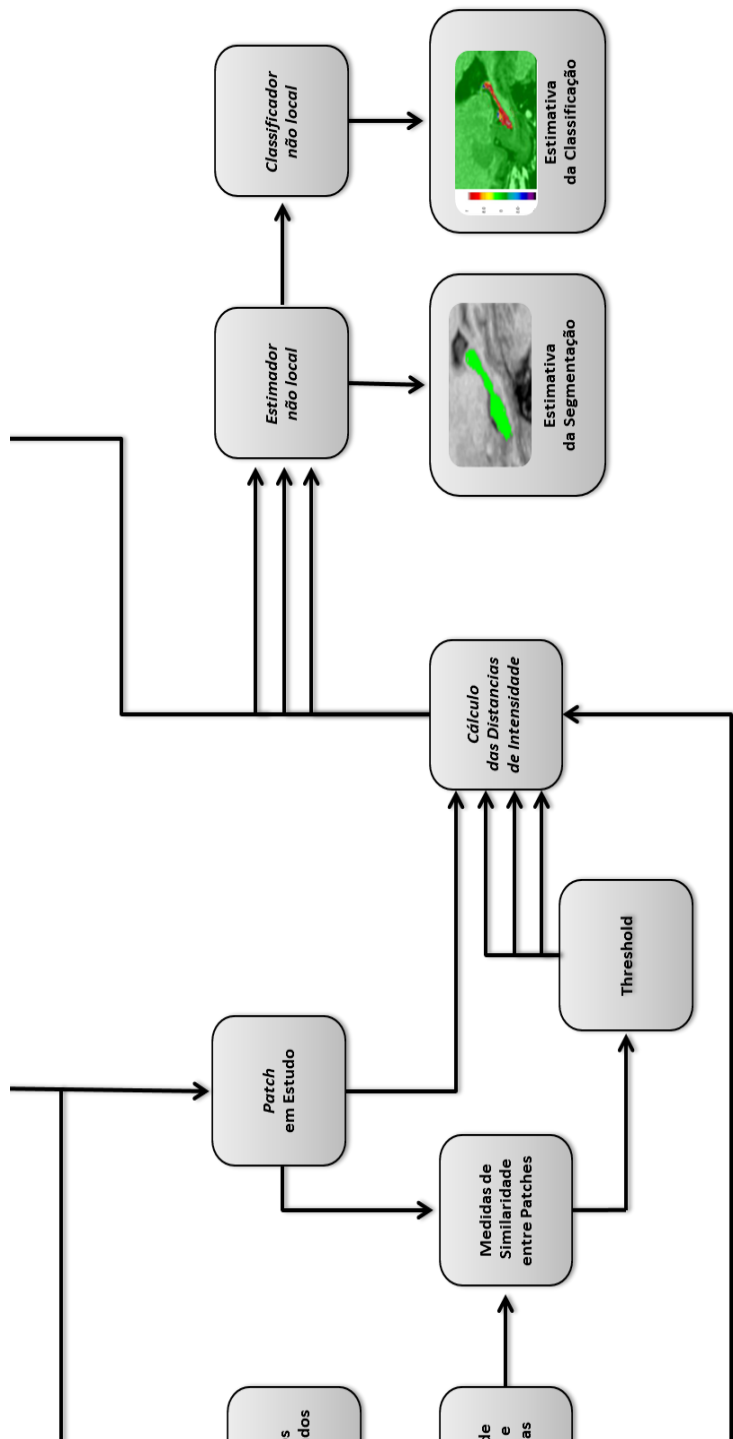


Figura 6.6- Diagrama representativo da implementação do algoritmo SNIPE baseado no filtro Laplaciano.

A primeira diferença está na forma como a base de dados é construída. Para além das imagens e máscaras dos sujeitos das populações de AD e CN, a biblioteca de treino é $y = -1\pi\sigma w^4 1 - x^2 + y^2 2\sigma w^2 e^{-(x^2 + y^2) 2\sigma w^2}$. (6.1), ao longo do volume da máscara inicial. Como a biblioteca é constante, estes mapas são pré-calculados, evitando computação adicional no algoritmo. No entanto, memória adicional é necessária, uma vez que os hyperslabs com os mapas de bordas dos N sujeitos selecionados são guardados num vetor para futura comparação.

O mapa das bordas da imagem do estudo é calculado no início de cada execução do *software*. Como se trata do processamento da porção de uma única imagem este passo não significa um aumento significativo no tempo de execução geral.

Com os mapas de bordas da biblioteca de treino e da imagem de estudo calculados, o algoritmo prossegue como anteriormente até ao cálculo das distâncias, onde a informação das bordas é utilizada. Este passo envolve o dobro da computação exigida pelo SNIPE, aumentando expressivamente o tempo de execução do método.

6.5. Esquema de validação

De forma a reduzir as diferenças com o SNIPE original, o mesmo tipo de classificação e esquema de validação foram usados nesta dissertação. Assim, uma análise discriminante linear foi usada como classificador e uma abordagem LOOCV foi usada para classificar todas as 818 imagens da base de dados escolhida. Estes foram usados com as implementações disponibilizadas pelo scikit-learn.

6.6. Sumário

Nenhuma versão do algoritmo SNIPE se encontra disponível publicamente, pelo que, uma das principais tarefas da dissertação foi a implementação da totalidade do método.

Todo o pré-processamento necessário foi realizado utilizando o pacote de ferramentas MINC. A construção do modelo de população específica, derivado da base de dados ADNI, requer um *cluster* com grande poder computacional, algo que não esteve disponível para esta dissertação. Assim, o atlas ICBM152 foi usado como modelo no registo.

Na primeira fase do estudo ADNI, existem 229 sujeitos de controlo e os 193 pacientes de Alzheimer. Deste conjunto, 239 máscaras hipocampais foram disponibilizadas. As imagens destes sujeitos foram integradas na biblioteca de treino. Para além destas, as 396 imagens restantes de sujeitos com MCI foram usadas na validação da implementação.

A linguagem escolhida para a implementação do algoritmo foi o C++. Esta linguagem é umas das linguagens mais utilizadas na atualidade, destacando-se pela sua excelente performance. Para evitar conversões de formato das imagens resultados do pré-processamento, a biblioteca de C++ do MINC foi escolhida para leitura das imagens.

A modificação do peso do algoritmo SNIPE não implica uma grande alteração da implementação do método. Para além das imagens e máscaras dos sujeitos das populações de AD e CN, a nova biblioteca de treino é construída com mapas de bordas calculados com filtro de Laplace. Depois disto, o algoritmo prossegue como anteriormente até ao cálculo das distâncias, onde a informação das bordas é utilizada.

Capítulo 7 Resultados e Discussão

Neste capítulo começa-se por referir as métricas utilizadas na avaliação da segmentação e dos classificadores, bem como testes estatísticos usados para comparar os dados. Posteriormente são expostos os resultados obtidos na segmentação e são feitos estudos com os volumes e as gradações agrupadas por população. É ainda feita uma avaliação visual qualitativa dos mapas de gradação obtidos. A qualidade dos biomarcadores é então avaliada pelos resultados dos respetivos classificadores. Por fim, uma discussão dos tempos de computação dos métodos é também efetuada.

7.1. Métricas de avaliação

No caso do SNIPE, vários estudos podem ser efetuadas para avaliar a sua eficácia. Primeiramente, uma vez que o volume do hipocampo é usado como biomarcador, a qualidade da segmentação deve ser quantificada. O coeficiente Dice é uma das várias medidas da sobreposição espacial entre duas imagens binárias. Este é normalmente usado para reportar a performance de segmentações, dando um maior peso aos casos onde as imagens coincidem. O seu valor varia entre 0, quando não existe nenhuma sobreposição, e 1, quando as duas imagens são completamente coincidentes. Sendo R o resultado da segmentação a ser avaliada e V a segmentação manual considerada como a delineação

$R \cap V)R+V$ (7.1) [54]:

$$Dice = \frac{2(R \cap V)}{|R|+|V|} \quad (7.1)$$

Alternativamente, o coeficiente Dice pode ser escrito em termos de verdadeiros positivos (VP), os vóxeis corretamente classificados, falsos positivos (FP), os vóxeis classificados erradamente como pertencente à estrutura, e falsos negativos (FN), o número

$R \cap V)R+V$ (7.1) é equivalente à $Dice = \frac{2 \times VP}{VP+FP+VP+FN}$. (7.2)

[101]:

$$Dice = \frac{2 \times VP}{(VP+FP)+(VP+FN)}. \quad (7.2)$$

Os métodos de CAD, tal como a maioria dos testes de diagnóstico, são normalmente avaliados em termos de sensibilidade, especificidade e exatidão. Similarmente ao Dice, os

seus valores variam entre 0 e 1, e são diretamente proporcionais qualidade do método. Estas medidas podem ser calculadas com os verdadeiros positivos (VP), o número correto de diagnósticos da doença, os verdadeiros negativos (VN), a quantidade testes onde a ausência de patologia é corretamente classificada, os falsos positivos (FP), os casos onde o teste indica, erradamente, a presença de doença e os falsos negativos (FN) onde patologias existentes não são detetadas. A sensibilidade (Sen), especificidade (Esp) e a exatidão (Exa) $Sen = \frac{TP}{TP+FN}$, (7.3), $Sen = \frac{TP}{TP+FN}$, (7.3)4 e $Sen = \frac{TP}{TP+FN}$, (7.3)5 [102].

$$Sen = \frac{TP}{TP+FN} \quad (7.3)$$

$$Esp = \frac{TN}{TN+FP} \quad (7.4)$$

$$Exa = \frac{TP+TN}{TP+FN+FP+TN} \quad (7.5)$$

Por fim, três testes estatísticos são usados. A regressão linear é usada para avaliar a variação dos biomarcadores com a idade. O teste McNemar examina mudanças em dados relacionados, sendo usado nesta dissertação para avaliar a diferença das distribuições dos biomarcadores entre as diferentes populações [103]. O teste *Wilcoxon-Mann-Whitney* compara a média de duas populações independentes, sem assumir que estas possuem uma distribuição normal [103]. Por isso, este foi usado para comparar os resultados das classificações dos métodos propostos.

7.2. Avaliação da segmentação do hipocampo

A segmentação da estrutura estudada pelo SNIPE é um passo extremamente importante do método. Para além do volume total ser usado como biomarcador, a gradação é apenas efetuada ao longo da segmentação estimada [66]. A Tabela 7.1 apresenta o valor médio de Dice das segmentações obtidas pela implementação do SNIPE. Este foi calculado por comparação com as 239 imagens cujas segmentações manuais foram fornecidas pelo ADNI.

Tabela 7.1- Valor médio do coeficiente Dice mais o desvio-padrão ($Dice \pm \sigma$) das segmentações do SNIPE implementado. Os resultados são provenientes dos 137 sujeitos CN e 102 sujeitos AD cujas segmentações hipocampais foram fornecidas pelo ADNI

Populações	Dice $\pm \sigma$
CN	0,780 \pm 0,071
AD	0,778 \pm 0,035
Total	0.779 \pm 0,058

Os valores da Tabela 5.1 indicam que o método obtém segmentações de qualidade semelhante, independentemente do estado clínico do sujeito. Esta característica é importante para um método que propõe o uso do volume como biomarcador.

O resultado médio de Dice reportado pelo SNIPE é, com os parâmetros ótimos, de 0,884 [28]. Este é consideravelmente superior ao resultado alcançado nesta dissertação. No entanto, este valor é relativo a sujeitos jovens e saudáveis, onde a variabilidade do hipocampo é bastante inferior.

A qualidade da segmentação do SNIPE baseado no filtro de Laplace foi avaliada da mesma forma. Os resultados obtidos estão presentes na Tabela 7.2.

Tabela 7.2- Valor médio do coeficiente Dice mais o desvio-padrão ($Dice \pm \sigma$) das segmentações do SNIPE-LoG. Os resultados são provenientes dos 137 sujeitos CN e 102 sujeitos AD cujas segmentações hipocampais foram fornecidas pelo ADNI

Populações	Dice $\pm \sigma$
CN	0,755 \pm 0,015
AD	0,813 \pm 0,057
Total	0,779 \pm 0,063

Quando comparado com os valores da Tabela 7.1, não existem grandes diferenças com a introdução da informação das bordas na segmentação. Embora um ligeiro aumento tenha sido verificada na segmentação dos sujeitos AD, este é obtido à custa de um

decréscimo do Dice médio dos sujeitos CN. Globalmente, a qualidade das duas abordagens é semelhante.

Embora o Dice médio permita comparar facilmente eficácias de segmentação entre métodos, análises qualitativas das segmentações resultantes ajudam a identificar em que circunstâncias o algoritmo tende a falhar. As imagens seguintes são apresentadas na vista sagital, por esta ser a orientação onde o hipocampo é visível mais facilmente. A Figura 7.1- Slice sagital da pior segmentação obtida pelo SNIPE. apresenta o resultado da pior segmentação obtida pelo SNIPE.

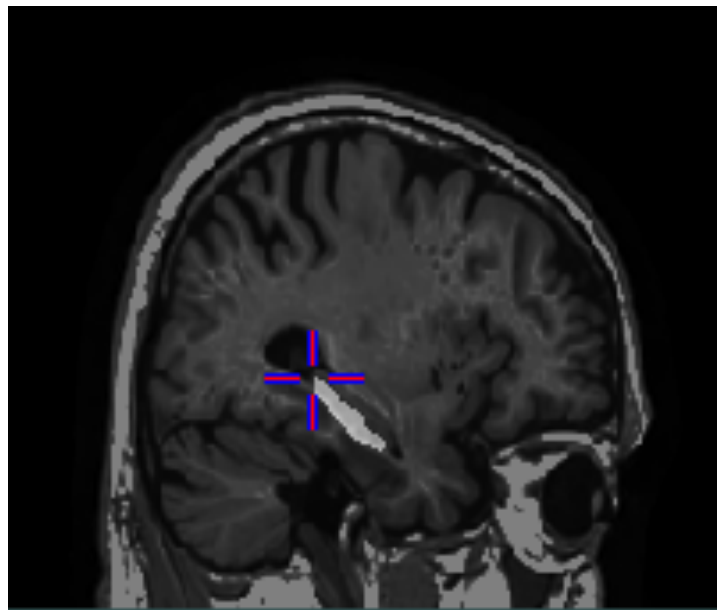


Figura 7.1- Slice sagital da pior segmentação obtida pelo SNIPE.

Neste sujeito, uma grande parte dorsal do hipocampo não é segmentada e, para além disso, toda a região ventral é erradamente incluída. Apesar destes erros óbvios, uma parte considerável do hipocampo foi corretamente identificada. No entanto, o Dice desta segmentação é extremamente baixo, de apenas 0,119. Um coeficiente desta ordem indica que a sobreposição entre a segmentação e a máscara real é muito baixa. Algo semelhante acontece na segmentação com SNIPE-LoG do mesmo sujeito, apresentada na Figura 7.2.

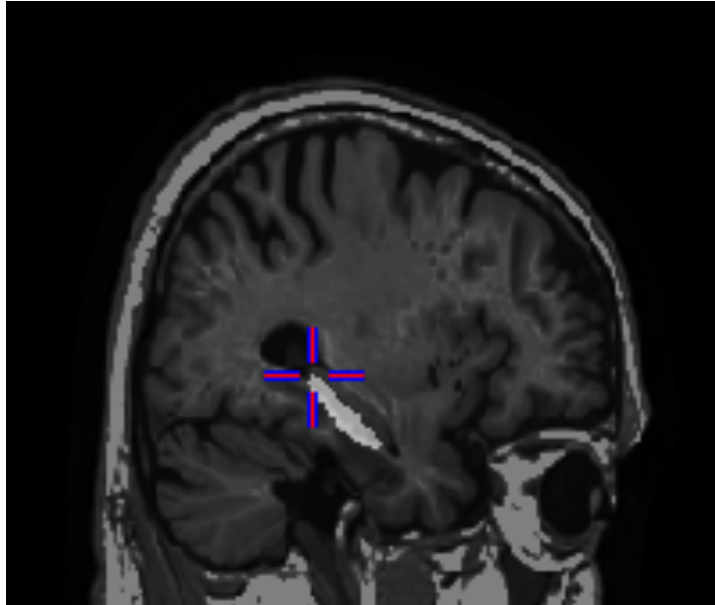


Figura 7.2- *Slice* sagital da pior segmentação obtida pelo SNIPE baseado no operado de Laplace.

Os mesmos problemas de segmentação da Figura 7.1 estão presentes nesta imagem. Existe uma sobrestimação ventral do hipocampo e grande parte da região dorsal do hipocampo não é identificada. De igual forma, o coeficiente Dice é muito baixo, de apenas 0,119. Para perceber os coeficientes obtidos, a máscara ADNI do sujeito foi também visualizada. Esta pode ser vista na Figura 7.3.

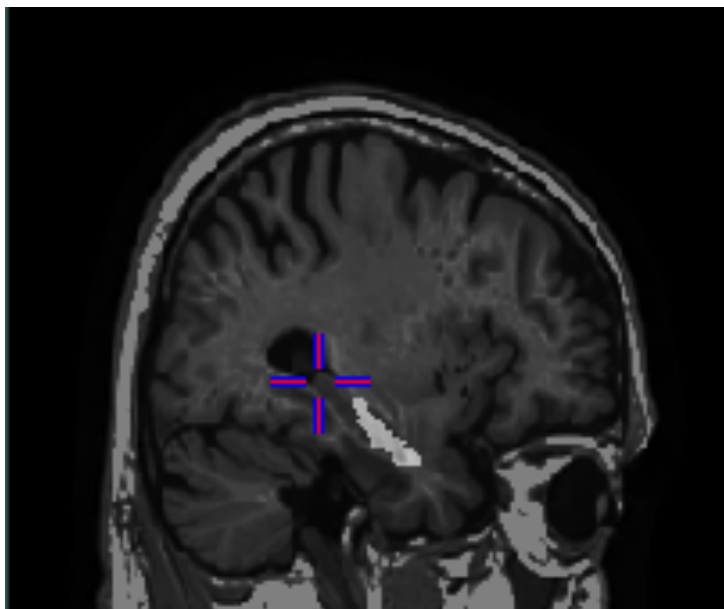


Figura 7.3- *Slice* sagital da máscara do hipocampo fornecida pelo ADNI.

Quando a máscara ADNI é sobreposta à imagem do sujeito correspondente, torna-se claro que não existe uma correta segmentação hipocampal. Neste *slice*, o hipocampo

encontra-se, nitidamente, numa zona mais caudal à região identificada. Para além disso, a própria forma da máscara não corresponde à geometria observada.

Nos ficheiros originais, uma sobreposição completa entre o hipocampo do sujeito e a zona delineada é verificada. Assim, a distorção observada na máscara da Figura 7.3 é o resultado do registo para o espaço estereotáxico efetuado no pré-processamento.

Como já foi referido, a cada uma das máscaras ADNI foi aplicada a mesma transformação que ao MRI correspondente. No entanto, existe uma ligeira diferença entre as resoluções dos ficheiros em que as máscaras e as imagens RM estão armazenadas e, embora esta não seja perceptível visualmente, é o suficiente para causar transformações espaciais diferentes. Embora a qualidade da transformação tenha sido avaliada qualitativamente para algumas imagens, o elevado número de MRIs impossibilitou a verificação de todas as 239 transformações. Assim, este erro foi apenas detetado depois de todos os testes terem sido efetuados. A implementação de Coupé et al. não sofreu do mesmo problema, uma vez que as imagens foram segmentadas manualmente no espaço estereotáxico [66].

A presença de informação errada nas máscaras ADNI tem grandes implicações nos resultados do algoritmo SNIPE. Primeiramente, estas foram usadas como base para a propagação dos *labels* para os restantes sujeitos AD e CN da base de dados. Os erros destas máscaras podem ser propagados para as restantes segmentações o que, ultimamente resulta na adição de mais erros na biblioteca de treino. Para além disso, dados de treino errados resultam, inevitavelmente, em erros de segmentação. Uma vez que o SNIPE propõe o uso do volume do hipocampo como biomarcador, isto pode diminuir a capacidade de diagnóstico do método. As deformações das segmentações também põem em causa a veracidade dos valores de Dice reportados nas Tabela 7.1 e Tabela 7.2. Por fim, as deformações das máscaras hipocampais podem ainda afetar o valor da gradação atribuído pelo SNIPE. Embora a gradação use *labels* representativos dos estados clínicos dos sujeitos na comparação, a região do mapa de gradação é definida pelo passo da segmentação. Erros de segmentação podem resultar na gradação de regiões cerebrais menos descritivas da doença.

Como já referido, os erros nas máscaras foram apenas detetados depois da realização de todos os testes e, devido ao tempo necessário para processar um número tão elevado de imagens, novas computações não puderam ser realizadas. Independentemente das possíveis irregularidades, todas as imagens da base de dados foram processadas pelo SNIPE e pelo SNIPE-LoG, tendo os seus biomarcadores sido extraídos e estudados.

7.3. Estudo volumétrico

Tal como em [60], um estudo volumétrico foi realizado, onde os volumes obtidos foram traçados em função da idade do sujeito para as quatro populações estudadas. A Figura 7.4 apresenta os resultados volumétricos obtidos pelo SNIPE implementado.

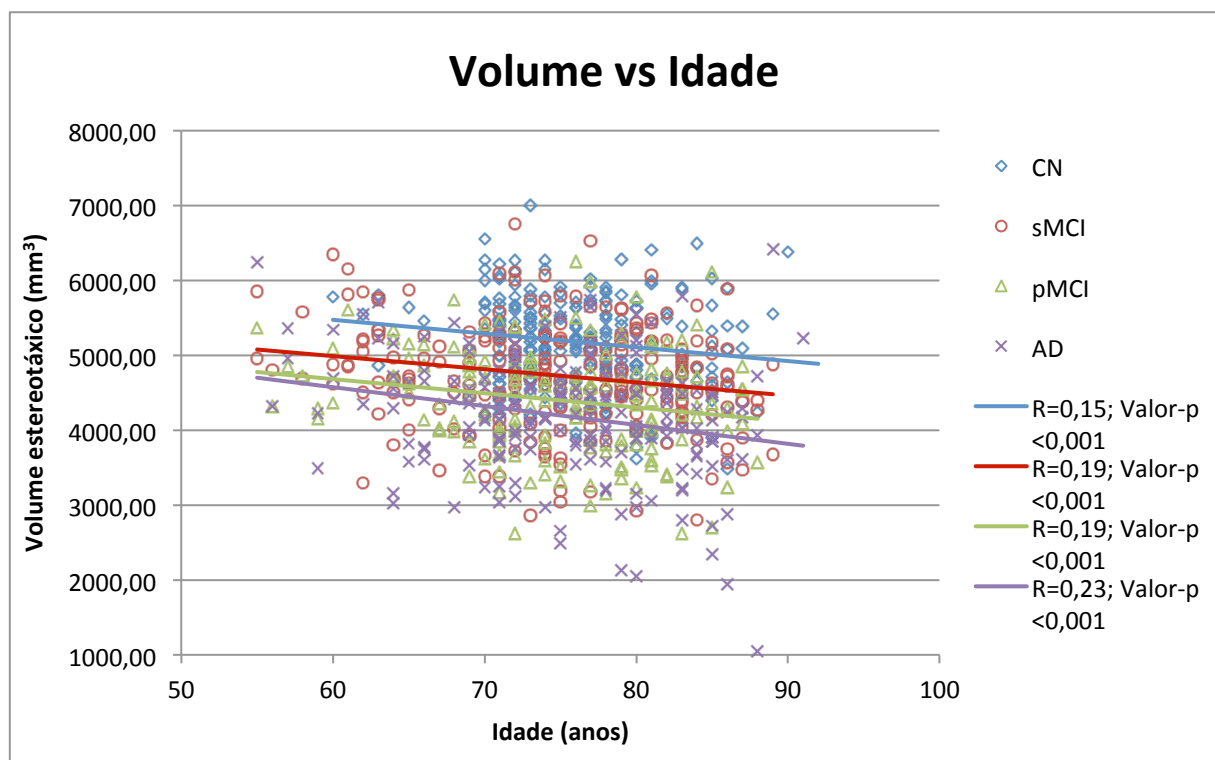


Figura 7.4- Estimativa pelo SNIPE implementado do volume do hipocampo em função da idade dos sujeitos. As regressões lineares são apresentadas para uma melhor visualização das tendências globais. O coeficiente de Pearson (R) e os valores-p das regressões lineares são apresentados na legenda.

Uma redução do volume do hipocampo com a idade é observada em todas as populações, sendo a sua significância estatística provada pelos baixos valores-p das regressões. Uma maior redução é notória na população AD, algo que pode ser explicado pela

adição da patologia à atrofia relacionada com a idade. Além disso, para a mesma idade, as correlações mostram uma tendência de separação dos volumes entre as populações.

As distribuições dos volumes para cada população são ainda apresentadas num diagrama de caixas na Figura 7.5.

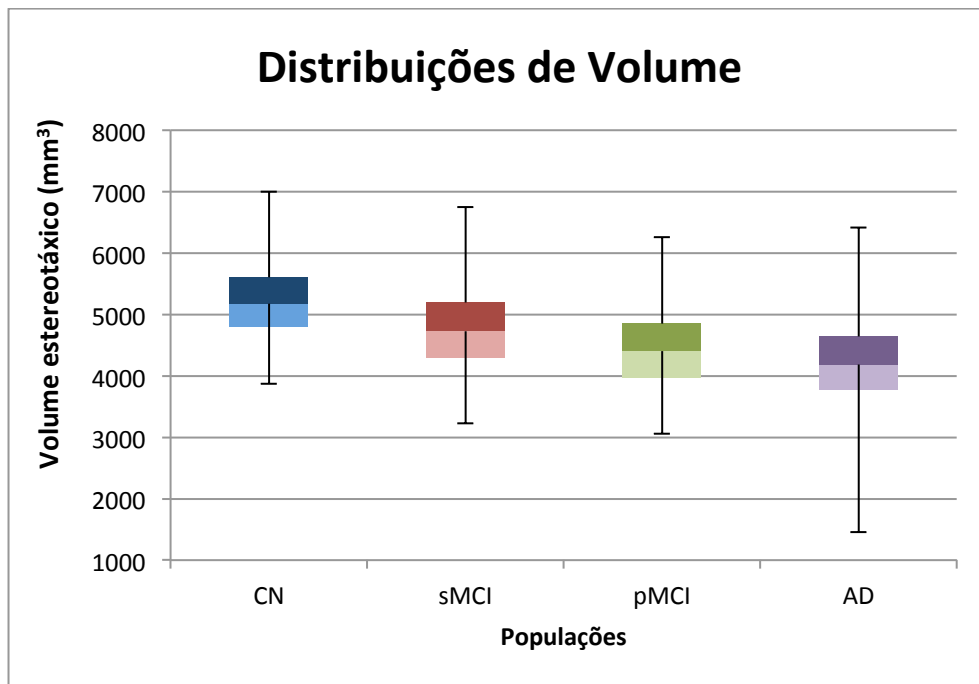


Figura 7.5- Diagrama de caixas das distribuições dos volumes, obtidos pela implementação do SNIPE implementado, para cada população.

De acordo com o teste de comparação Wilcoxon-Mann-Whitney, todas populações têm médias de volume significativamente diferentes (valor-p <0,05). Todos os valores-p estão apresentados na Tabela 1 do Anexo II – Resultados dos testes estatísticos. A ordem dos volumes é a esperada ($AD < pMCI < sMCI < CN$), sendo estes inversamente proporcionais à gravidade dos estados clínicos. É ainda verificada uma maior variação na distribuição do volume da população AD. Esta pode ser explicada por variações causadas pela doença, ou por uma maior dificuldade do método em segmentar hipocampus com maior grau de atrofia e, consequentemente, com maior grau de variabilidade. Por fim, é ainda visível maior separação entre os volumes CN e sMCI do que entre os de AD e pMCI.

Todas as conclusões deste estudo volumétrico foram também verificadas por Coupé et al. [60]. No entanto, a grandeza dos volumes reportados é bastante superior em [60]. Isto sugere que os erros de registo das máscaras ADNI afetaram, significativamente, a

capacidade do método de segmentar o hipocampo. Porém, o facto das relações dos volumes entre as populações se manterem, aponta para que as dificuldades na segmentação sejam semelhantes para todas as populações.

O mesmo estudo volumétrico foi efetuado com o SNIPE-LoG. Os resultados são visíveis na Figura 7.66 e na Figura 7.7.

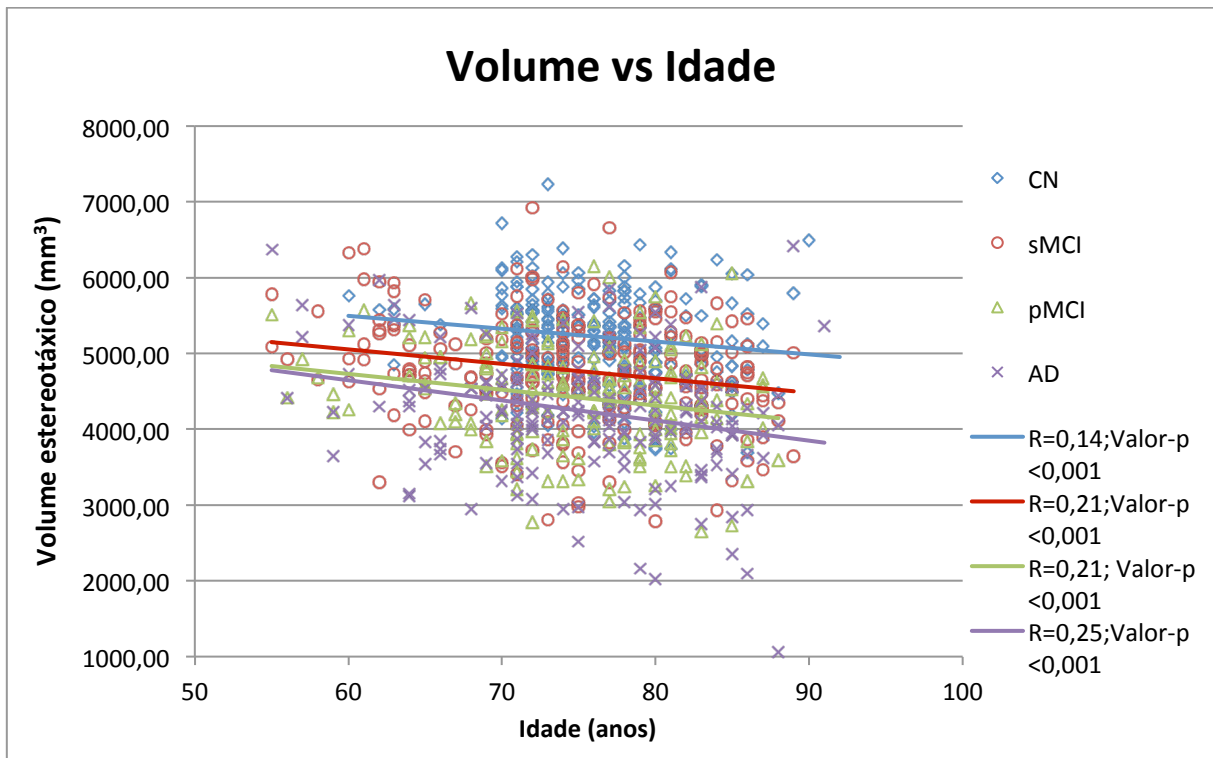


Figura 7.6- Estimativa pelo SNIPE-LoG do volume do hipocampo em função da idade dos sujeitos. As regressões lineares são apresentadas para uma melhor visualização das tendências globais. O coeficiente de Pearson (R) e os valores-P das regressões lineares são apresentados na legenda.

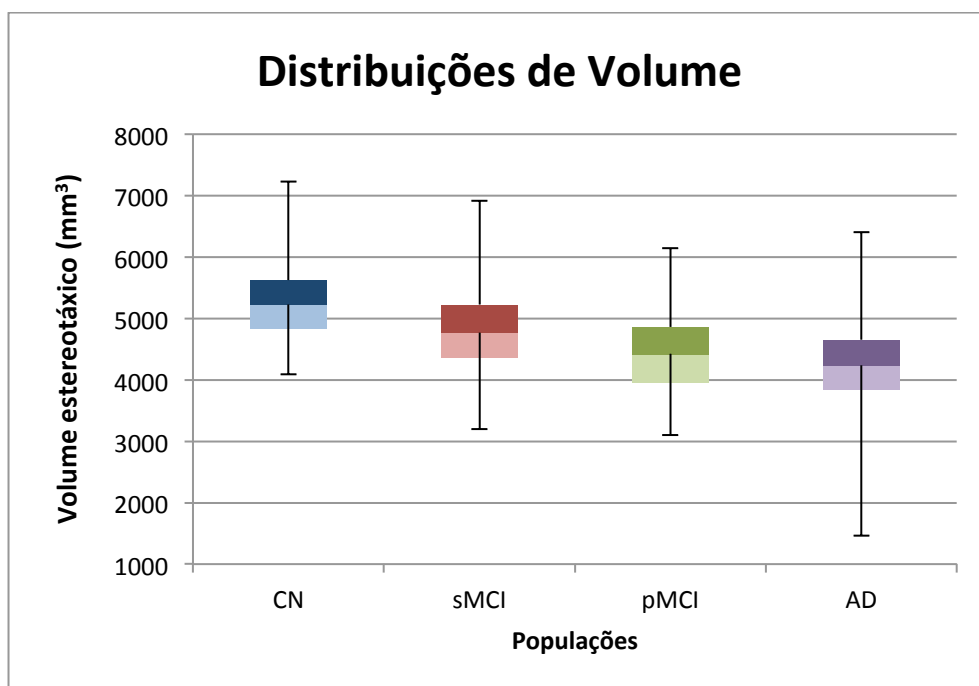


Figura 7.7- Diagrama de caixas das distribuições dos volumes, obtidos pelo SNIPE-LoG, para cada população.

De uma forma geral, as distribuições dos volumes são semelhantes às anteriores. De igual forma, o teste Wilcoxon-Mann-Whitney mostrou que as distribuições dos volumes mantêm-se significativamente diferentes entre cada par de populações (valor-p <0,05). Todos os valores-p estão apresentados na Tabela 2 do Anexo II – Resultados dos testes estatísticos. Assim, todas as observações efetuadas para o SNIPE são válidas para o SNIPE-LoG.

Os valores dos volumes obtidos continuam a ser inferiores aos reportados por Coupé et al., no entanto, quando comparado com os resultados da implementação do SNIPE, um ligeiro incremento médio é observado em cada população. Os problemas com as máscaras na biblioteca de treino fazem com que, apesar dos Dices obtidos serem semelhantes, a eficiência de segmentação do SNIPE e do SNIPE-LoG não seja necessariamente a mesma. Logo, este ligeiro aumento na ordem dos volumes obtidos pelo SNIPE-LoG pode sugerir uma melhor capacidade de segmentação.

Como referido em Base de dados ADNI16.2. , desde da publicação do trabalho de Coupé et al., uma nova fase do projeto ADNI foi iniciada, sendo que novos diagnósticos foram disponibilizados para alguns dos sujeitos. Assim, desde de 2012, 30 sujeitos transitaram de MCI para AD. A influência destas conversões foi estudada através das

comparações das distribuições dos volumes obtidos pelo SNIFE e pelo SNIFE-LoG. Estas estão apresentadas na Figura 7.8 e na Figura 7.9, respetivamente.

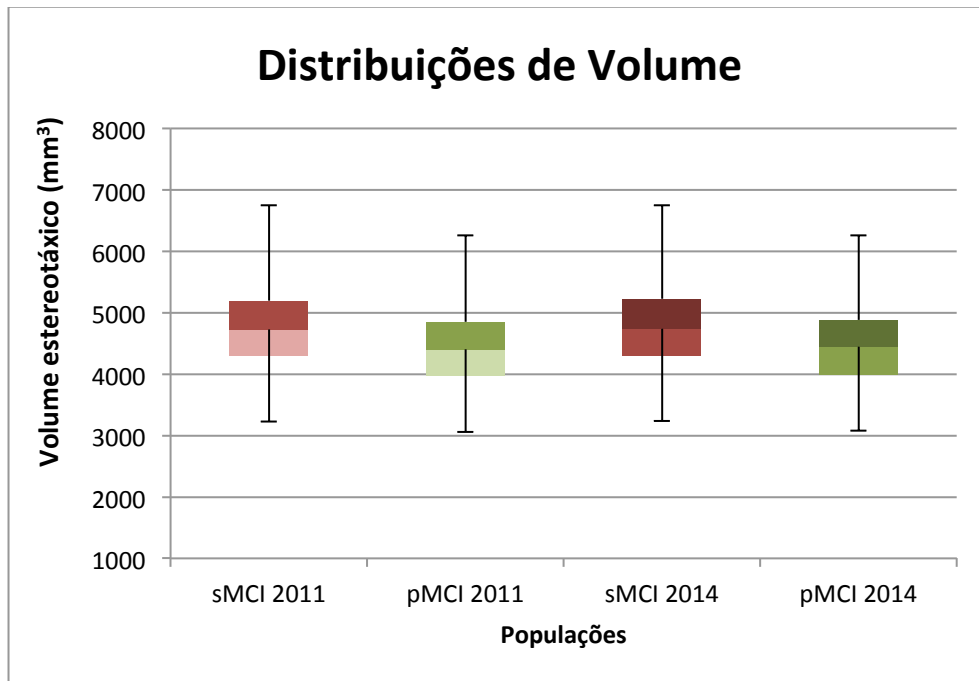


Figura 7.8- Diagramas de caixas das distribuições dos volumes, obtidos pelo SNIFE implementado, para as populações sMCI e pMCI obtidas com a informação de conversão até 2011 ou 2014.

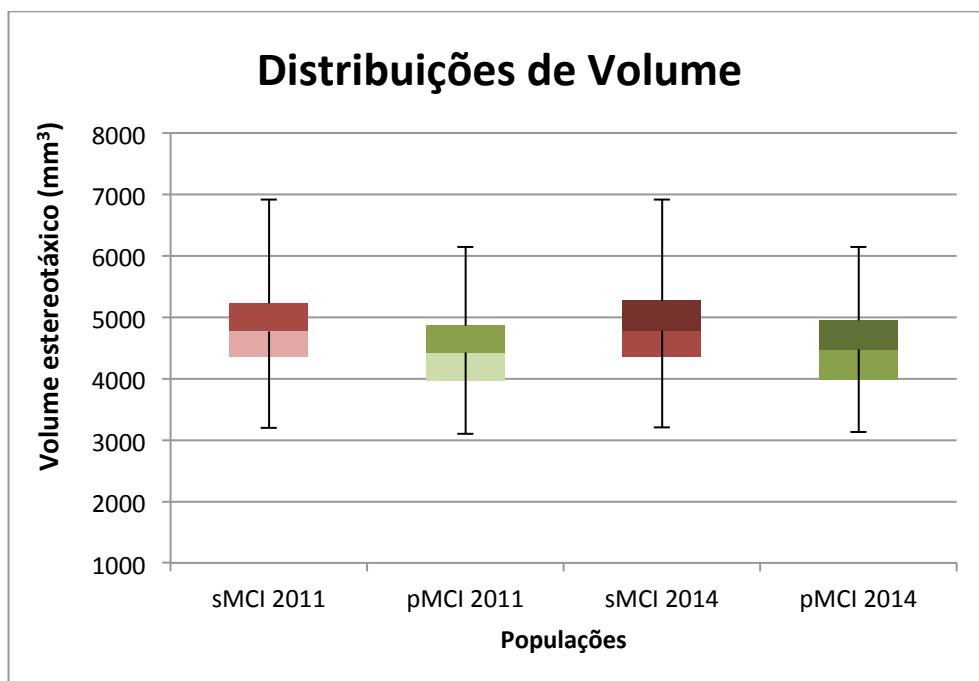


Figura 7.9- Diagramas de caixas das distribuições dos volumes, obtidos pelo SNIFE-LoG, para as populações sMCI e pMCI obtidas com a informação de conversão até 2011 ou 2014.

As Figura 7.8 e a Figura 7.9 mostram que não existem grandes alterações na distribuição do volume da população sMCI criada em 2011 (sMCI 2011) e a construída em 2014 (sMCI 2014), ou entre população pMCI de 2011 (pMCI 2011) e a de 2014 (pMCI 2011), independentemente do método usado. O teste Wilcoxon-Mann-Whitney foi ainda usado para comparar os volumes das populações construídas nos dois pontos temporais. Os valores-p resultantes foram, para os dois algoritmos, superiores a 0,05, pelo que uma diferença significativa entre as populações não foi comprovada. Todos os valores-p estão apresentados na Tabela 12 e na Tabela 4 do Anexo II – Resultados dos testes estatísticos. A proximidade dos dois estados clínicos pode justificar que a introdução errada dos 30 volumes, em populações com mais de 150 sujeitos, não seja suficiente para alterar a distribuição global dos volumes das populações.

7.4. Estudo da gradação

A Figura 7.10 apresenta os valores da gradação médios obtidos pela implementação do SNIPE.

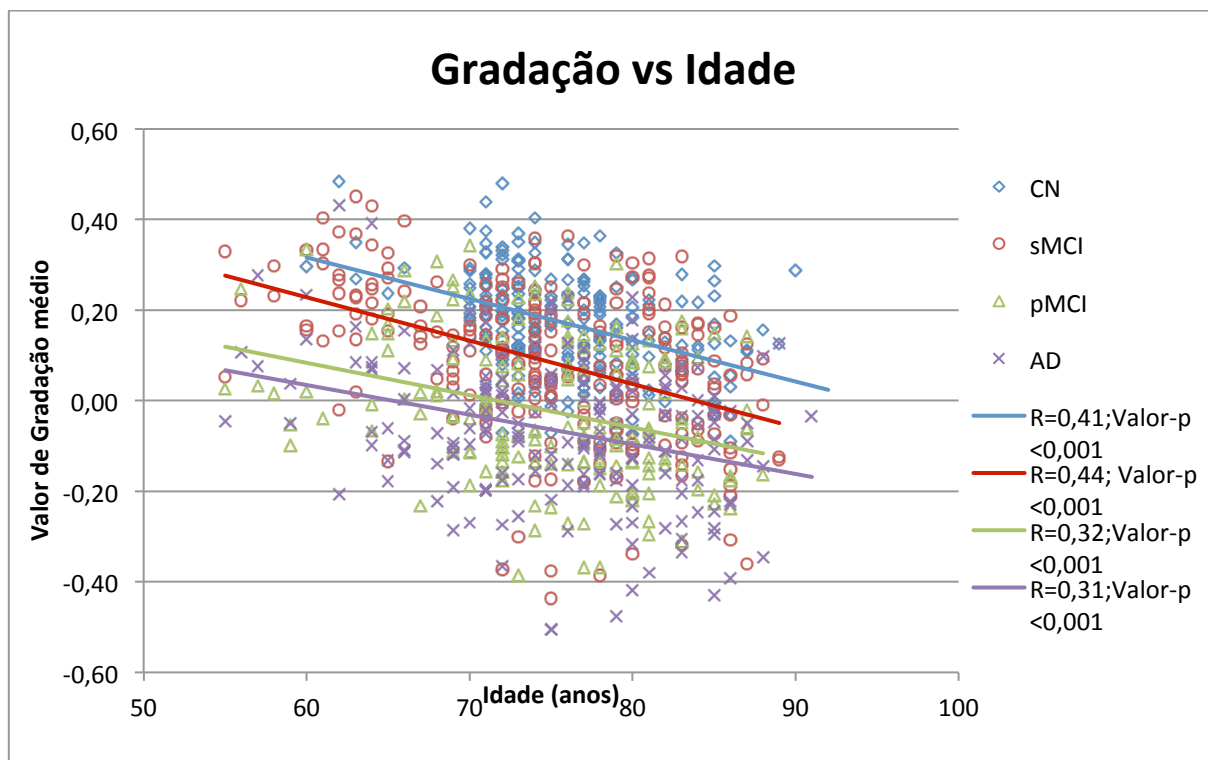


Figura 7.10- Estimativa pelo SNIPE implementado da gradação média do hipocampo em função da idade dos sujeitos. As regressões lineares são apresentadas para uma melhor visualização das tendências globais. O coeficiente de Pearson (R) e os valores-P das regressões são apresentados na legenda.

Para o hipocampo, os valores da gradação de todas as populações são estatisticamente correlacionados com a idade (valores-p <0,05), diminuindo quando esta aumenta. Para além disso, quando comparado com o estudo volumétrico, os coeficientes de correlação obtidos pela gradação são superiores. Como comprovado na posterior comparação dos biomarcadores, a maior correlação com a idade permite uma melhor distinção entre diferenças anatómicas relacionadas com a idade e alterações relacionadas com a patologia.

Como no estudo volumétrico, as distribuições das gradações médias são apresentadas para cada população em diagramas de caixas. Estes são visíveis na Figura 7.11.

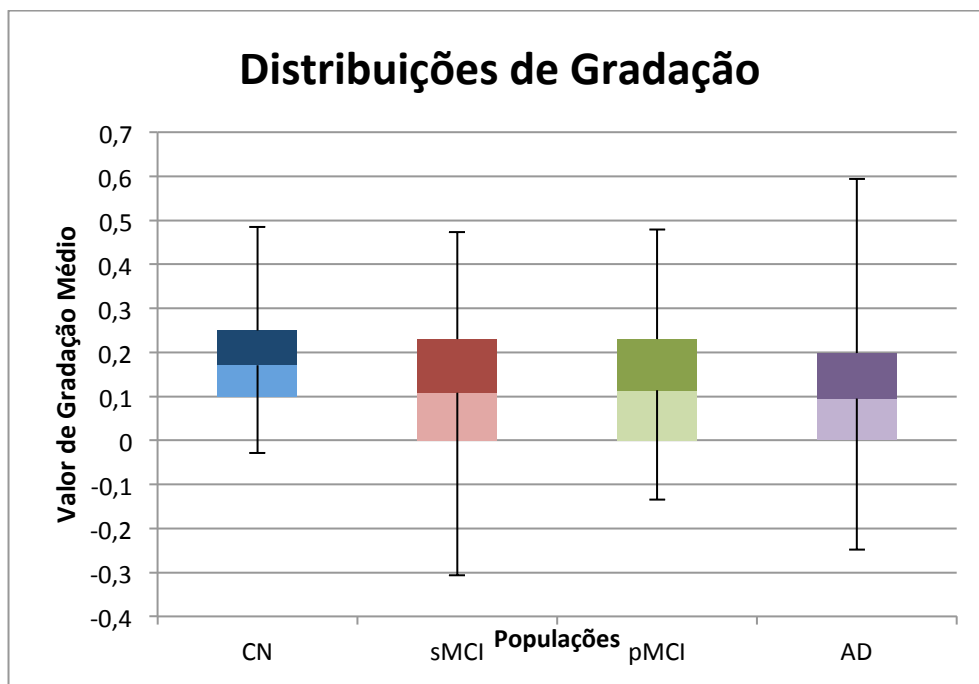


Figura 7.11- Diagrama de caixas das distribuições das gradações médias, obtidas pela implementação do SNIFE, para cada população.

Como esperado, os sujeitos CN obtiveram os maiores valores da gradação, em contraste com os pacientes com AD, que obtiveram os mais baixos. O teste Wilcoxon-Mann-Whitney mostrou que as distribuições têm médias significativamente diferentes (valor- $p < 0,05$) entre cada par de populações. Todos os valores- p estão apresentados na Tabela 11 do Anexo II – Resultados dos testes estatísticos

Embora a média das populações siga a ordem esperada ($AD < pMCI < sMCI < CN$), a separação das distribuições das populações é muito inferior à obtida por Coupé et al. em [60]. Mais uma vez, as diferenças podem ser justificadas pelos problemas das máscaras hipocampais usadas na base de dados. A gradação de cada vóxel é obtida por uma pesagem de *labels* expressivos do estado clínico dos *patches* de treino. Diretamente, este cálculo não é afetado pelas máscaras de treino presentes na biblioteca, no entanto, o biomarcador é obtido pela gradação média da estrutura estudada. Se a qualidade da segmentação do hipocampo for inferior, a quantidade de tecido cerebral patológico avaliado é menor. No caso dos estados clínicos muito próximos, como o sMCI e pMCI, a perda da informação, presente no tecido do hipocampo que não é segmentado, resulta numa drástica queda da diferenciação entre populações. A proximidade das distribuições das quatro populações

indica que que segmentações obtidas não foram suficientes para obter mapas da gradação semelhantes aos obtidos por Coupé et al..

O mesmo estudo da gradação foi efetuado com o SNIPE-LoG. Os resultados são visíveis na Figura 7.12 e Figura 7.13.

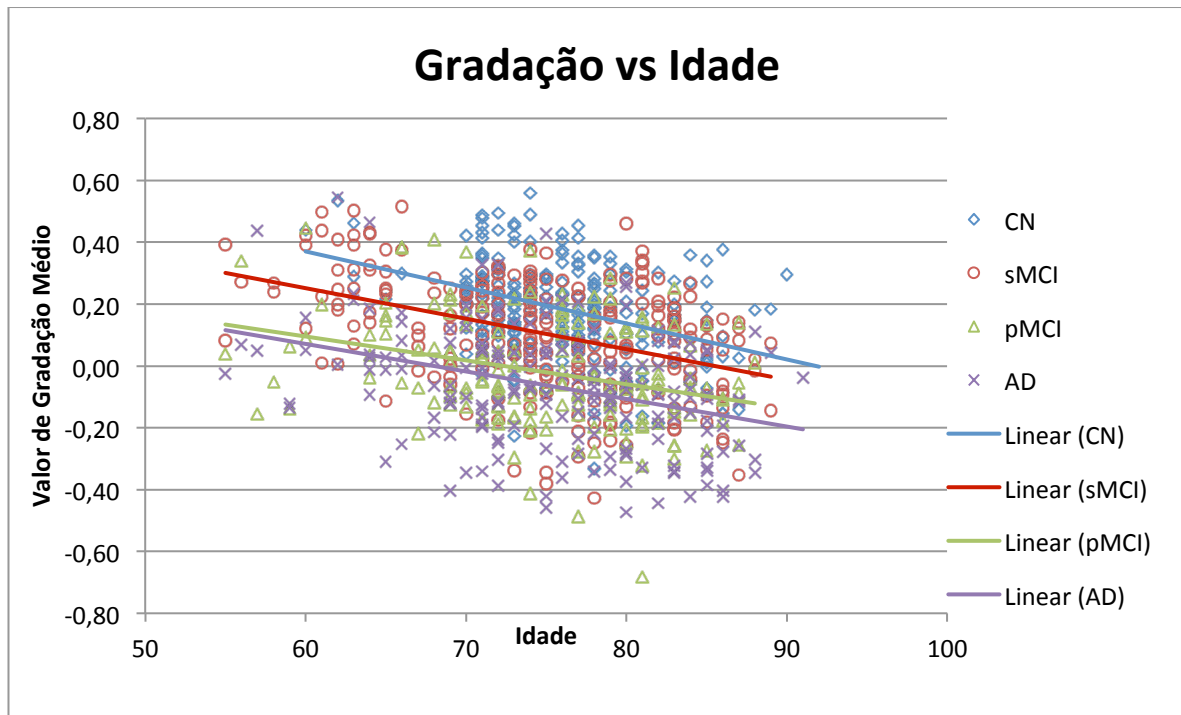


Figura 7.12- Estimativa pelo SNIPE-LoG da gradação média do hipocampo em função da idade dos sujeitos. As regressões lineares são apresentadas para uma melhor visualização das tendências globais. O coeficiente de Pearson (R) e os valores-P das regressões são apresentados na legenda.

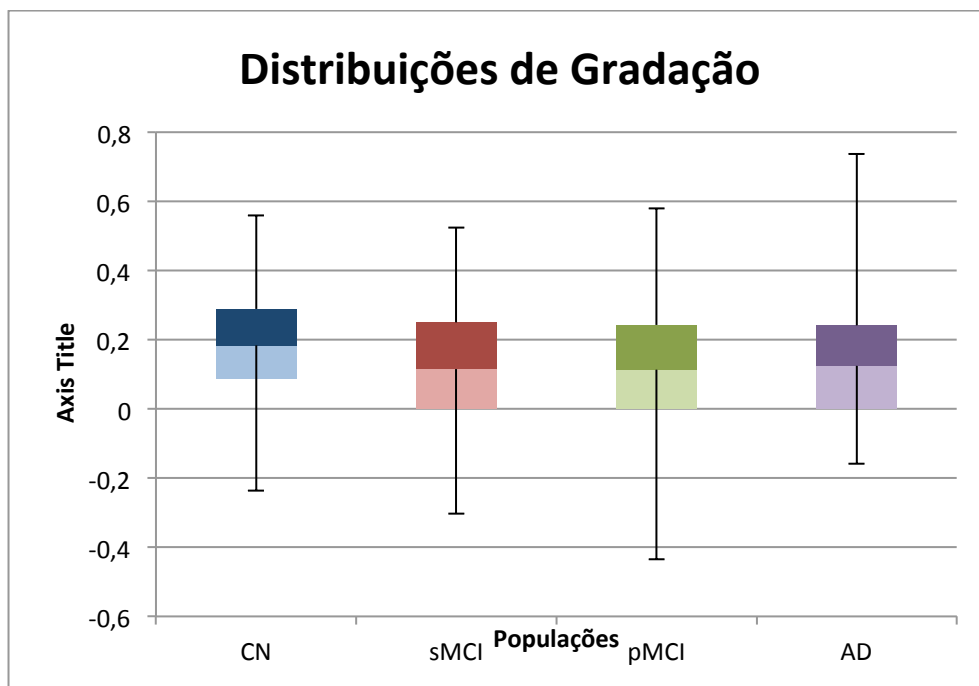


Figura 7.13- Diagrama de caixas das distribuições da gradação média, obtidos pelo SNIPE-LoG, para cada população.

O teste *Wilcoxon-Mann-Whitney* provou que todas as populações mantêm médias de gradação significativamente diferentes entre si (valor-p <0,05). Todos os valores-p estão apresentados na Tabela 13 do Anexo II – *Resultados dos testes estatísticos*. Globalmente, não existe grande diferença com as distribuições obtidas pela implementação SNIPE, pelo que, a análise anterior também é válida para o SNIPE-LoG. Embora estes resultados não sejam promissores, estes não provam, irrefutavelmente, a utilidade do filtro Laplaciano no cálculo da gradação. Numa abordagem por patches não locais, os primeiros locais a ser afetados por defeitos de segmentação são as zonas de maior variabilidade, como por exemplo, as fronteiras das estruturas. Logo estas são as primeiras zonas a ser afetados pela existência de máscaras deformadas na base de dados. O principal objetivo do uso do SNIPE-LoG é a melhoria da comparação nas zonas da fronteira do hipocampo, se estas não são incluídas na estimativa da segmentação, é natural que a introdução da informação das arestas na gradação não provoque grandes alterações no resultado.

A influência da conversão de 30 pacientes para AD desde de 2012 nas distribuições da gradação média foi também estudada para a gradação. Para isso, a Figura 7.14 e a Figura

7.15 mostram, respetivamente, as distribuições das populações sMCI e pMCI de 2012 e de 2014 obtidas pelo método SIPE e pelo SNIPE-LoG.

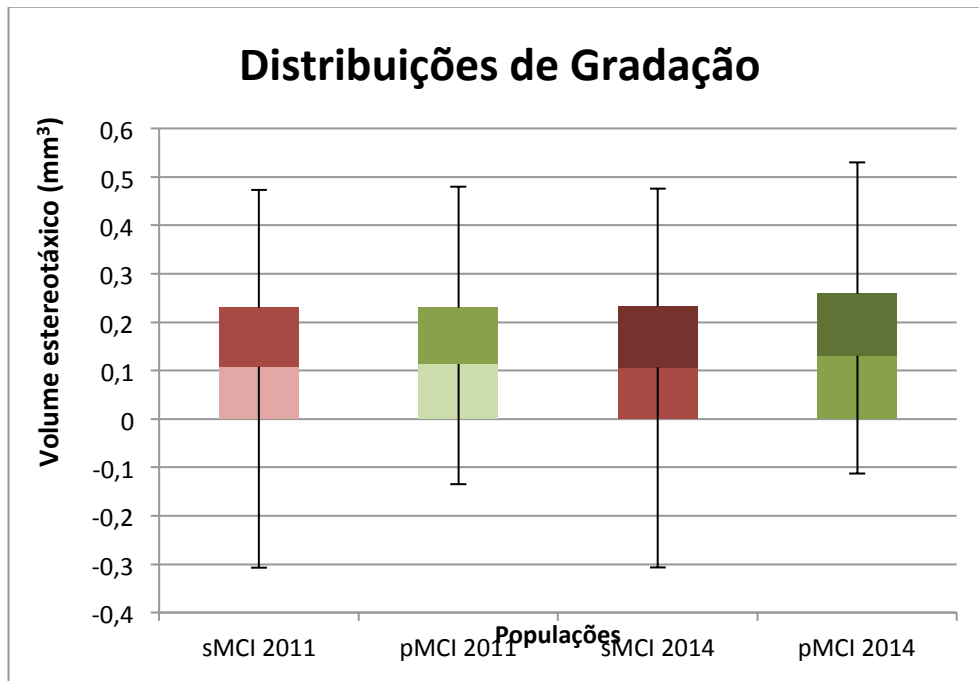


Figura 7.14- Diagramas de caixas das distribuições da gradação média, obtidos pelo SNIPE, para as populações sMCI e pMCI obtidas com a informação de conversão até 2011 ou 2014.

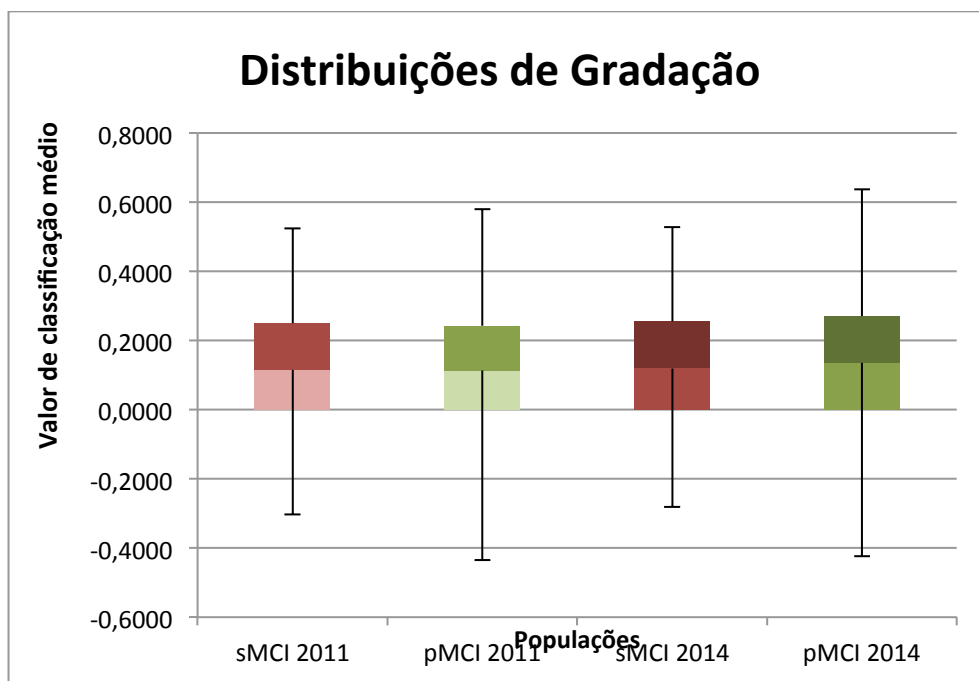


Figura 7.15- Diagramas de caixas das distribuições da gradação média, obtidos pelo SNIPE-LoG, para as populações sMCI e pMCI obtidas com a informação de conversão até 2011 ou 2014

Tal como com o volume, nenhum dos métodos aparenta detetar alterações nas distribuições das populações. Quando sujeitos ao teste *Wilcoxon-Mann-Whitney*, os valores da gradação dos mesmos tipos de populações, construídas em 2012 e 2014, obtiveram valores-p superiores a 0,05, pelo que uma diferença significativa da média de valores não é assegurada. Todos os valores-p estão apresentados na Tabela 12 e na Tabela 4 do Anexo II – *Resultados dos testes estatísticos*.

Novamente, a proximidade dos estados clínicos pode fazer com que, o diagnóstico errado de 30 pacientes não seja o suficiente para provocar alterações nas distribuições de populações com mais de 150 sujeitos.

Finalmente, uma avaliação visual na mudança de gradação com a idade entre as populações é apresentada na Figura 7.16.

Para os dois métodos, a gradação estimada em pacientes AD é visualmente inferior à dos sujeitos CN. Esta tendência pode também ser observada entre as populações sMCI e pMCI. Para além disso, uma diminuição global nos mapas de gradação com a idade é também notória.

A avaliação visual permite ainda uma distinção nos resultados dos mapas de gradação fornecidos pelo SNIPE e pelo SNIPE-LoG. Analisando o sujeito CN mais idoso da Figura 7.16, por exemplo, a região mais dorsal da cauda hipocampal tem valores claramente mais expressivos ($|g(x_i)|$ mais próximos de 1) no resultado SNIPE-LoG do que no SNIPE. O mesmo acontece na zona mais ventral da cabeça do hipocampo do sujeito AD mais idoso. Isto sugere que, em zonas de grandes fronteiras do hipocampo, a introdução da informação das bordas é responsável por uma melhor estimação da gradação. Estas diferenças são inesperadas, uma vez que as distribuições de gradação obtidas pelos dois métodos não apresentam grandes diferenças. Isto pode dever-se pelo facto do biomarcador usado ser obtido pela média dos mapas de gradação. Uma vez que as zonas de fronteira têm volumes inferiores ao resto do hipocampo, as suas estimativas acabam por assumir uma menor importância no valor final do biomarcador. Uma solução possível seria estudar novas formas de obter biomarcadores a partir dos mapas de gradação, como por exemplo, usar a média da gradação de apenas as zonas próximas nas bordas do hipocampo. Isto permitiria,

simultaneamente, acelerar a execução do algoritmo, uma vez que evitaria a necessidade de segmentar todo o hipocampo.

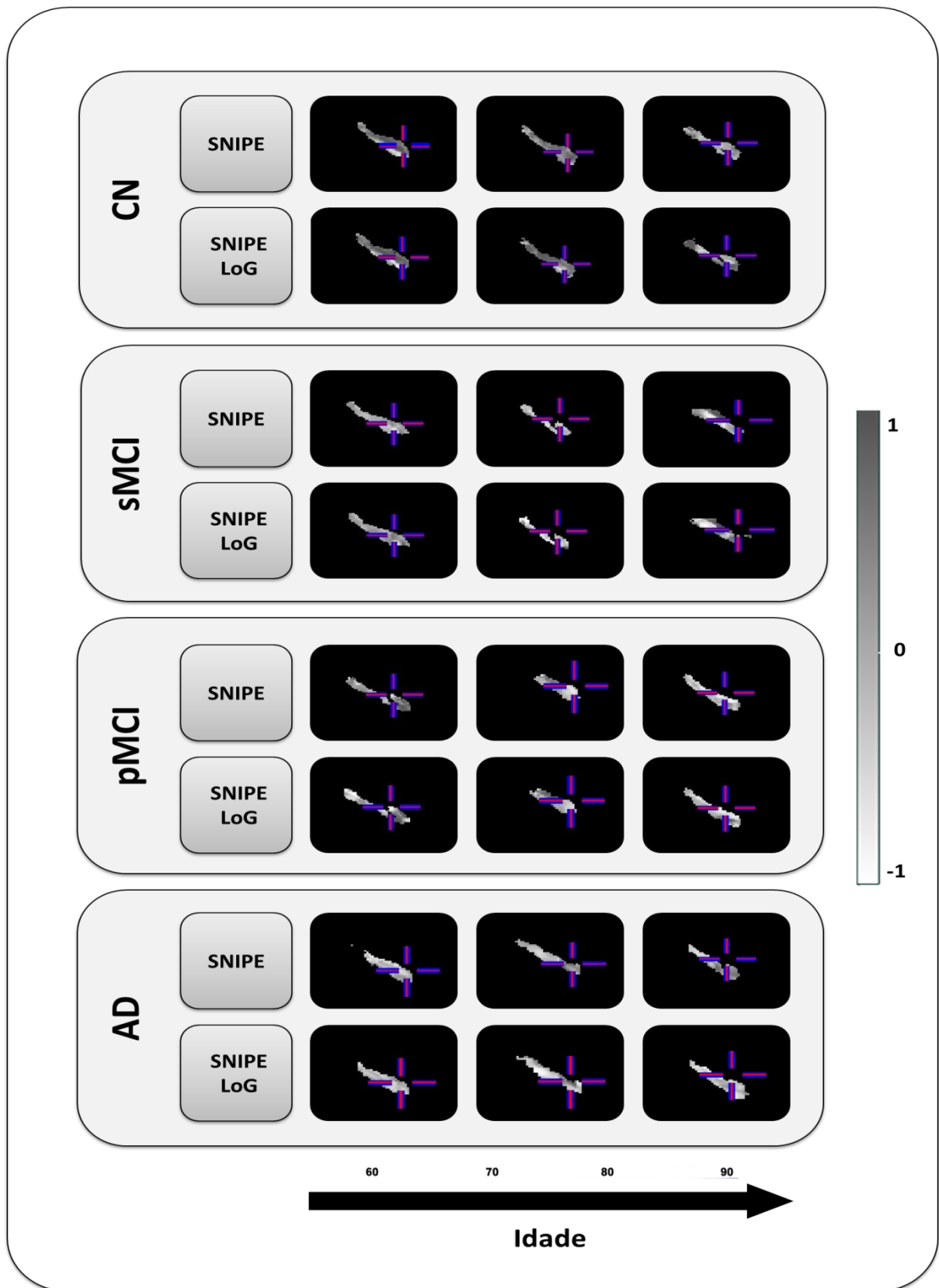


Figura 7.16- Exemplos de mapa de gradação de acordo com a idade e com a população.

7.5. Comparação dos biomarcadores do SNIPE

Uma avaliação das taxas de sucesso da classificação é essencial para qualquer método de diagnóstico. A Tabela 7.3 apresenta os valores dos classificadores LDA, com uma abordagem LOOCV, obtidos pelos biomarcadores imagiológicos do SNIPE e do SNIPE-LoG.

Tabela 7.3- Resultados obtidos para a diferenciação das diferentes populações com os biomarcadores obtidos do hipocampo. São apresentados simultaneamente os resultados obtidos com a implementação do SNIPE e com o SNIPE-LoG. Estes foram obtidos usando a análise discriminante linear com um procedimento LOOCV. Os valores na tabela correspondem à exatidão (Exa), sensibilidade (Sen) e a especificidade (Esp) em percentagem (%) e o valor-p (p) do teste Mcnemar para avaliar a performance dos testes comparados com uma classificação aleatória. É ainda apresentado o Valor-p do teste Mcnemar para avaliar significância da diferença dos resultados obtidos pelos dois métodos

	Exa%	Sen%	Esp%	Valor-p	Exa %	Sen%	Esp%	Valor-p	
AD vs. CN	SNIPE				SNIPE-LoG				SNIPE vs. SNIPE-LoG
Volume	78	73	83	p<0.0001	79	73	84	p<0.0001	p=0,332
Gradação	84	76	91	p<0.0001	80	74	86	p<0.0001	p=0,051
Volume+ Gradação	85	77	91	p<0.0001	84	81	87	p<0.0001	p=0,877
pMCI vs. CN	SNIPE				SNIPE-LoG				SNIPE vs. SNIPE-LoG
Volume	72	59	82	p<0.0001	76	65	83	p<0.0001	p=0,004
Gradação	81	67	91	p<0.0001	78	68	86	p<0.0001	p=0,203
Volume+ Gradação	80	69	88	p<0.0001	79	73	84	p<0.0001	p=0,728
AD vs. sMCI	SNIPE				SNIPE-LoG				SNIPE vs. SNIPE-LoG
Volume	66	52	77	p<0.0001	67	52	79	p<0.0001	p=0,606
Gradação	67	58	75	p<0.0001	68	62	74	p<0.0001	p=0,653
Volume+ Gradação	70	60	77	p<0.0001	70	63	75	p<0.0001	p=1,000
pMCI vs. sMCI	SNIPE				SNIPE-LoG				SNIPE vs. SNIPE-LoG
Volume	61	32	82	p<0.0001	63	35	82	p<0.0001	p=0,211
Gradação	70	53	82	p<0.0001	67	50	79	p<0.0001	p=0,124
Volume+ Gradação	71	58	80	p<0.0001	70	60	77	p<0.0001	p=0,885
AD vs. pMCI	SNIPE				SNIPE-LoG				SNIPE vs. SNIPE-LoG
Volume	55	76	30	p=0,1712	56	81	26	p=0,1070	p=0,628
Gradação	56	75	33	p=0,0963	54	76	29	p=0,2000	p=0,560
Volume+ Gradação	54	73	33	p=0,2000	53	74	28	p=0,4100	p=0,327
sMCI vs. CN	SNIPE				SNIPE-LoG				SNIPE vs. SNIPE-LoG

									LoG
Volume	65	62	69	p<0.0001	65	63	66	p<0.0001	p=0,831
Gradação	64	62	65	p<0.0001	61	60	62	p<0.0001	p=0,267
Volume+ Gradação	65	62	68	p<0.0001	64	61	66	p<0.0001	p=0,542

Pelos valores da Tabela 7.3, à exceção do AD vs. pMCI, os biomarcadores do dois métodos têm um desempenho significativamente melhor ($p < 0,05$) do que uma classificação aleatória. Tal como esperado, as exatidões das classificações diminuem quando populações com estados patológicos mais próximos são comparadas. Deste modo, o melhor resultado foi encontrado para a comparação dos dois extremos clínicos, AD vs. CN. Semelhantemente, uma grande diferença foi encontrada entre as populações pMCI e CN. Embora a mesma ordem de resultados seja esperada para a comparação AD vs. sMCI, valores menos discriminativos foram encontrados. Estas classificações sugerem que a população total MCI é mais próxima dos sujeitos AD do que dos sujeitos CN. Alternativamente, o argumento pode ser feito para a existência de heterogeneidade no grupo sMCI, onde alguns sujeitos podem ainda converter para AD, aproxima-o da população AD. Estas observações são suportadas pelas classificações sMCI vs. CN e AD vs. pMCI, onde uma classificação muito superior é conseguida em sMCI vs. CN.

Quando comparado com os resultados reportados por Coupé et al., diferenças consideráveis são encontradas. De uma forma geral, todas as classificações da Tabela 7.3 são ligeiramente inferiores aos apresentados em 5.6. Porém, no caso da separação de populações muito semelhantes, como AD vs. pMCI e sMCI vs. CN, reduções drásticas das exatidões são verificadas. No caso do pMCI vs. sMCI, ainda que a exatidão seja semelhante à publicada por Coupé et al., a sensibilidade obtida é muito inferior. Nestes casos, as maiores diferenças estão nas taxas de sucesso que envolvem a gradação. Isto sugere que a classificação baseada na gradação é mais sensível a alterações no algoritmo do que a baseada no volume.

Como referido em 7.2. , a maior diferença com implementação do SNIPE por Coupé et al. foi a utilização de máscaras no espaço original da imagem, sendo estas posteriormente registadas com a transformação aplicada à imagem correspondente. Devido às pequenas

diferenças de resolução nas máscaras e nas imagens originais, a sobreposição completa das máscaras resultantes com o hipocampo não é garantida.

Os resultados do estudo volumétrico sugeriram que, apesar da presença de informação falsa na biblioteca de treino afetar a capacidade de segmentação do algoritmo, as relações volumétricas entre populações são aproximadamente semelhantes. A proximidade das classificações baseadas no volume com as obtidas por Coupé et al., para a maioria das comparações, suporta esta hipótese.

As exatidões obtidas na classificação com a gradação vão de encontro ao que já foi referido no estudo da gradação. No caso de comparações de populações clinicamente distintas, a segmentação obtida foi suficiente para obter resultados da gradação próximos dos publicados por Coupé et al.. À medida que a semelhança entre as populações testadas aumenta, a perda da informação, presente no tecido do hipocampo que não é segmentado, resulta numa drástica queda da exatidão do classificador.

Os resultados do SNIPE-LoG são muito semelhantes aos obtidos pela implementação SNIPE. Embora as sensibilidades sejam, de uma forma geral, ligeiramente superiores, estas são obtidas com um decaimento da especificidade. Isto leva a valores de exatidão muito próximos entre os dois métodos, não tendo sido provada uma diferença significativa para a maioria das classificações. A única garantia de diferença significativa é encontrada na distinção entre pMCI e CN ($p < 0,05$), usando apenas o volume como biomarcador, onde o SNIPE-LoG conseguiu os melhores resultados.

Como já foi referido no estudo volumétrico, a ordem dos volumes obtidos pelo SNIPE-LoG aproximam-se mais aos reportados por Coupé et al., o que sugere uma melhor capacidade de segmentação. No entanto, à exceção do pMCI vs. CN, não existem melhorias nas classificações baseadas no volume. Isto pode dever-se a uma de duas razões. Os maiores volumes obtidos pelo SNIPE-LoG podem dever-se a um aumento dos falsos positivos, não havendo uma correlação com o tamanho real da estrutura. Alternativamente, o volume pode ser um biomarcador pouco sensível a alterações, e as diferenças nos volumes obtidos não são suficientes para causar mudanças nas classificações.

Tal como no estudo da gradação, as classificações com as gradações médias do SNIPE-LoG não obtiveram grandes diferenças com a implementação do SNIPE. Como já foi

referido, numa abordagem por *patches* não locais, os primeiros locais a ser afetados por defeitos de segmentação são as zonas de maior variabilidade. Por isso, se as fronteiras do hipocampo não são incluídas na segmentação devido à presença de máscaras incorretas na base de dados, é natural que a introdução da informação das arestas no método não resulte numa alteração significativa dos biomarcadores obtidos. Para além disso, alterações na gradação das fronteiras podem acabar por ser diluídas no cálculo da média, uma vez que as zonas de fronteira têm volumes inferiores ao resto do hipocampo. Finalmente, os melhores parâmetros para a aplicação do SNIPE-LoG podem não ter sido usados. O desvio padrão do *kernel* do filtro Laplaciano usado foi o sugerido por Bhjile et al., sendo este usado empiricamente, e para a filtragem de imagens inteiras do cérebro. Um estudo da influência deste valor nos biomarcadores permitiria saber se estes poderiam ser melhorados. Além disso, um estudo do efeito do tamanho dos *patches* e do volume de procura é aconselhável, de forma a determinar se os parâmetros ótimos do SNIPE-LoG são os mesmos do SNIPE. Estes estudos não foram realizados nesta dissertação devido ao tempo de computação requerido para a realização de todo o processamento necessário.

Por fim, as taxas de sucesso dos dois métodos, utilizando as populações construídas em 2014, são apresentadas na Tabela 7.4. Como a única diferença com as populações de 2011 é a passagem de 30 sujeitos de sMCI para pMCI, apenas classificações com estas populações são apresentadas.

Tabela 7.4- Resultados obtidos para a diferenciação das diferentes populações, criadas com os dados divulgados até 2014, com os biomarcadores obtidos do hipocampo. São apresentados simultaneamente os resultados obtidos com a implementação do SNIPE (SNIPE) e com o SNIPE-LoG (SNIPE-LoG). Estes foram obtidos usando a análise discriminante linear com um procedimento LOOCV. Os valores na tabela correspondem à exatidão (Exa),sensibilidade (Sen) e a especificidade (Esp) em percentagem (%) e o valor-p (p) do teste Mcnemar para avaliar a performance dos testes comparados com uma classificação aleatória. É ainda apresentado o Valor-p do teste Mcnemar para avaliar significância da diferença dos resultados obtidos pelos dois métodos

	Exa%	Sen%	Esp%	Valor-p	Exa %	Sen%	Esp%	Valor-p	
pMCI vs. CN	SNIPE				SNIPE-LoG				SNIPE vs. SNIPE-LoG
Volume	0,73	0,60	0,82	p<0.0001	0,76	0,65	0,83	p<0.0001	p=0,003
Classificação	0,81	0,68	0,91	p<0.0001	0,78	0,67	0,86	p<0.0001	p=0,385
Volume+	0,80	0,69	0,88	p<0.0001	0,79	0,73	0,84	p<0.0001	p=0,527

Classificação									
AD vs. sMCI	SNIPE				SNIPE-LoG				SNIPE vs. SNIPE-LoG
Volume	0,66	0,53	0,77	p<0.0001	0,67	0,53	0,78	p<0.0001	p=0,701
Classificação	0,68	0,59	0,75	p<0.0001	0,69	0,63	0,74	p<0.0001	p=0,670
Volume+ Classificação	0,69	0,60	0,77	p<0.0001	0,71	0,64	0,76	p<0.0001	p=0,981
pMCI vs. sMCI	SNIPE				SNIPE-LoG				SNIPE vs. SNIPE-LoG
Volume	0,59	0,51	0,66	p<0.0001	0,60	0,55	0,64	p<0.0001	p=0,201
Classificação	0,63	0,59	0,67	p<0.0001	0,61	0,59	0,62	p<0.0001	p=0,224
Volume+ Classificação	0,63	0,58	0,67	p<0.0001	0,64	0,64	0,65	p<0.0001	p=0,719
AD vs. pMCI	SNIPE				SNIPE-LoG				SNIPE vs. SNIPE-LoG
Volume	0,57	0,75	0,37	p=0,1531	0,56	0,79	0,30	p=0,1142	p=0,738
Classificação	0,55	0,72	0,35	p=0,0919	0,53	0,74	0,30	p=0,1934	p=0,476
Volume+ Classificação	0,54	0,70	0,36	p=0,2124	0,52	0,71	0,31	p=0,3974	p=0,237
sMCI vs. CN	SNIPE				SNIPE-LoG				SNIPE vs. SNIPE-LoG
Volume	0,65	0,62	0,69	p<0.0001	0,65	0,63	0,66	p<0.0001	p=0,791
Classificação	0,64	0,62	0,66	p<0.0001	0,61	0,59	0,63	p<0.0001	p=0,319
Volume+ Classificação	0,65	0,61	0,68	p<0.0001	0,64	0,61	0,67	p<0.0001	p=0,452

Quando comparado com os valores da Tabela 7.3, a maioria das classificações gera resultados iguais ou muito semelhantes às classificações com as populações de 2011. Isto sugere que a adição ou remoção de 30 sujeitos numa das populações não causa grandes alterações nos classificadores usados. Isto vai de acordo com as similaridades nas distribuições verificadas no estudo volumétrico e no estudo de gradação.

A maior alteração é verificada na classificação pMCI vs. sMCI. Como estas duas populações são as únicas a ser alteradas, é natural que esta seja a maior diferença. Neste caso, os dois métodos tiveram uma grande queda na especificidade, indicando que grande parte dos sujeitos que converteram para AD desde de 2012 estava a ser classificada como sMCI. Esta queda na especificidade resulta numa diminuição da exatidão do classificador quando comparado com as populações de 2012. É possível que na maioria destes 30 casos, devido ao facto da conversão para AD ser apenas verificada ao fim de vários anos, uma

atrofia significativa do hipocampo não estivesse visível na imagem de *screening* usada no teste.

7.6. Tempo de computação

Uma análise dos tempos de execução de todos os passos dos métodos propostos é apresentada. Os testes foram efetuados usando um processador Intel Core 2 Duo T6400 a 2,0GHz e 4 GB de RAM pelo que, todos os valores reportados são referentes ao uso deste *hardware*. Quando comparações com os tempos publicados por Coupé et al. forem efetuadas, é importante lembrar que um processador Intel Core 2 Quad Q6600 a 2,66GHz e 4 GB de RAM foi usado.

O primeiro passo do método consiste na aplicação do pré-processamento à imagem em estudo. Como referido em 6.1. , a biblioteca MINC foi usada para realizar todos os passos de pré-processamento. Dependendo da imagem, a filtragem por média não local e a correção das não-homogeneidades demoram cerca de 1 minuto cada, tempos similares aos reportados por Coupé et al. com as suas ferramentas em C, desenvolvidas internamente [28]. No entanto, o `standard_pipeline` usado para realizar os restantes passos do pré-processamento requer cerca de 45 minutos para ser completado. Embora uma parte significativa seja consumida pelo algoritmo BEaST, uma grande porção do tempo é ocupado no registo da imagem. Este é muito superior aos 2 minutos reportados por Coupé et al.. Assim, existe a uma oportunidade de melhoria na forma em como o pré-processamento é efetuado, podendo ser procurada uma biblioteca de processamento mais rápida e eficiente.

Relativamente à execução do SNIPE, uma redução no seu período de computação foi tentado com a implementação do algoritmo em C++. Os únicos tempos de execução reportados por Coupé et al. foram apresentados em [66], onde é mencionado que, para $N=20$, a segmentação e a gradação do hipocampo esquerdo e direito de um sujeito demora cerca de 4 minutos. Para realizar uma comparação justa, os tempos de execução dos métodos implementados foram calculados com o uso do mesmo número de sujeitos.

Para o SNIPE implementado, uma média de 3,5 minutos foram necessários para obter uma segmentação e um mapa de gradações para um sujeito. Como supracitado, um equipamento com características inferiores ao de Coupé et al. foi usado para os testes e,

mesmo assim, um tempo de computação inferior foi conseguido. Isto comprova a eficiência da implementação efetuada.

A implementação do SNIPE-LoG foi efetuada com o conhecimento *a priori* que este exigiria uma maior carga computacional. A única alteração do método é a introdução das diferenças das imagens das bordas no parâmetro de peso. Como, para além das diferenças de intensidades, estas diferenças necessitam de ser calculadas, o dobro do número de operações é necessário para o cálculo do peso de cada vóxel. De facto, para N=20, o SNIPE-LoG demora, em média, 4,8 minutos, necessitando cerca de 37% mais tempo para processar um sujeito.

O uso de um maior número de sujeitos aumenta significativamente o tempo de computação dos dois métodos. Para os resultados reportados, com N=100, a implementação do SNIPE requer cerca de 25 minutos enquanto o SNIPE-LoG necessita de 50 minutos por sujeito. Estes valores não podem ser comparados com a implementação de Coupé et al., uma vez que para este número de sujeitos, o tempo de computação não é publicado. O cálculo de pesos é o passo computacionalmente mais exigente deste tipo de abordagens e, realmente, quando o número de *patches* comparados aumenta (devido ao aumento do número de sujeitos comparados), o SNIPE-LoG tende a demorar o dobro do tempo do SNIPE. Uma vez que a alteração dos vóxeis da fronteira do hipocampo pode não ser suficiente para alterar significativamente o valor final da média da gradação, o aumento do tempo de computação pode não ser recompensador. No futuro, técnicas para aumentar a eficiência do método podem ser implementadas, como por exemplo, o uso da informação das arestas no método de pré-seleção de *patches*. Alternativamente, é ainda possível limitar o estudo apenas a zonas da imagem com grandes gradientes, evitando a necessidade de segmentar todo o hipocampo, aumentando a velocidade do método, e focando a gradação em zonas da fronteira hipocampal, onde a atrofia é mais expressiva.

7.7. Sumário

Para além dos estudos efetuados por Coupé et al., uma avaliação da segmentação do método foi realizada. Os Dice das segmentações obtidas foram calculados por comparação com as 239 imagens cujas segmentações manuais foram fornecidas pelo ADNI. Uma

diferença nas resoluções das máscaras e das imagens fez com que, a aplicação das transformações espaciais das imagens RM às máscaras resultasse nas suas distorções. O erro foi apenas detetado depois de todos os testes terem sido efetuados porque, embora a qualidade da transformação tenha sido avaliada qualitativamente para algumas imagens, o elevado número de imagens impossibilitou a verificação de todas as 239 transformações. A introdução de informação errada na biblioteca de treino reduz a qualidade das segmentações e das gradações estimadas. Além disso, estes erros comprometeram a veracidade dos Dices obtidos.

No estudo volumétrico foi verificada redução do volume do hipocampo com a idade em todas as populações. Além disso, para a mesma idade, as correlações mostram uma tendência de separação dos volumes entre as populações. A média dos volumes das populações seguiu a ordem esperada, AD < pMCI < sMCI < CN. Embora estas observações vão de acordo aos resultados de Coupé et al., a grandeza dos volumes obtidos é bastante inferior. Isto sugere que os erros de registo das máscaras ADNI afetaram, significativamente, a capacidade do método de segmentar o hipocampo.

No estudo de gradação, os valores de todas as populações são estatisticamente correlacionados com a idade, diminuindo quando esta aumenta. Embora a média das populações siga a ordem esperada, a separação das distribuições das populações é muito inferior à obtida por Coupé et al.. Mais uma vez, as diferenças podem ser justificadas pelos problemas das máscaras hipocampais usadas na base de dados.

Classificadores LDA, com uma abordagem LOOCV, foram usados para avaliar a taxa de sucesso de cada biomarcador. Os biomarcadores estudados tiveram, à exceção do AD vs. pMCI, um desempenho significativamente melhor do que uma classificação aleatória. O melhor resultado foi encontrado para a comparação dos dois extremos clínicos, AD vs. CN, havendo uma diminuição das exatidões das classificações à medida que populações com estados patológicos mais próximos são comparadas. Quando comparado com os resultados reportados por Coupé et al., as exatidões obtidas são ligeiramente inferiores, piorando à medida que comparações de populações mais próximas são efetuadas.

De uma forma geral, todos os resultados do SNIPE-LoG são muito próximos aos obtidos pela implementação do SNIPE. Embora as médias dos volumes segmentados tenham

sido ligeiramente superiores, esta não correspondeu a uma melhoria na classificação baseada no volume. No cálculo da gradação nenhuma diferença nos resultados é comprovada, sendo possível que as alterações na gradação das fronteiras acabem por ser diluídas no cálculo da média, uma vez que as zonas de fronteira têm volumes inferiores ao resto do hipocampo.

Por fim, tempos de computação inferiores aos reportados por Coupé foram conseguidos, o que comprova a eficiência da implementação efetuada.

Capítulo 8 Conclusão e trabalho futuro

Neste capítulo é sintetizado o trabalho desenvolvido e são expostas as conclusões mais relevantes. São ainda propostos, em trabalho futuro, alguns caminhos para melhorar o sistema desenvolvido.

8.1. Conclusão

O Alzheimer é o tipo mais comum de demência, afetando mais de 35.6 milhões de pessoas em todo o mundo. Embora nenhuma cura seja ainda conhecida, um tratamento precoce ajuda a atenuar os sintomas. Apesar disto, o diagnóstico tradicional envolve um processo longo e penoso para o paciente, sujeito a grande subjetividade por parte do médico e pressupõem um certo avanço da doença [1].

O principal objetivo da dissertação foi o desenvolvimento e implementação de um método completamente automático que permita auxiliar o diagnóstico de pacientes com Alzheimer. Para isso, foi implementado um algoritmo de segmentação e gradação automática do hipocampo usando o algoritmo de *patches* não locais propostos por Coupé [60]. Uma alteração ao método foi também proposta, onde se inclui informação das arestas das imagens na classificação efetuada.

Os resultados dos estudos apresentados demonstram que os biomarcadores calculados conseguem uma boa separação entre sujeitos CN e pacientes AD. A melhor classificação foi obtida com o uso do volume e da gradação média, com uma exatidão de 85%. Esta é muito próxima à exatidão de 87% publicada por Coupé [60]. No entanto, a utilização da implementação para prever as conversões para Alzheimer obteve resultados piores que os esperados. As limitações da implementação deveram-se a problemas encontrados nas máscaras do hipocampo usadas na biblioteca de treino.

O facto das segmentações utilizadas não estarem no mesmo espaço onde a análise é efetuada, provou ser a maior dificuldade do método. As transformações aplicadas, originadas dos registos das imagens MR correspondentes para o espaço estereotáxico, acabaram por deslocar ou deformar as máscaras relativamente ao hipocampo. A introdução de informação errada na biblioteca de treino limitou a qualidade dos biomarcadores obtidos,

especialmente no caso das classificações de populações clinicamente próximas. Além disso, as máscaras de hipocampo erradas põem em questão os Dice obtidos na avaliação da segmentação.

Os resultados do método SNIPE-LoG proposto não foram suficientes para provar melhorias significativas aos biomarcadores do SNIPE. Embora as médias dos volumes segmentados tenham sido ligeiramente superiores, os problemas das máscaras impossibilitaram comparações definitivas entre as segmentações dos dois métodos. Em termos de gradação média, grandes diferenças também não foram verificadas. Concluiu-se que o biomarcador de gradação sugerido pelo SNIPE não é o melhor para o SNIPE-LoG, uma vez que as zonas de fronteira têm volumes inferiores ao resto do hipocampo, e as estimativas destas zonas acabam por assumir uma menor importância no valor médio da gradação.

Os tempos de computação observados foram inferiores aos reportados por Coupé em [66], pelo que o uso de C++ permitiu uma implementação mais rápida e eficiente.

Em conclusão, embora os métodos implementados sejam suficientes para distinguir sujeitos CN com pacientes AD, a partir de uma única imagem de RM, a previsão de demência foi fortemente limitada pelos problemas originados nas máscaras do hipocampo usadas.

8.2. Trabalho Futuro

Dada a limitação de tempo para o desenvolvimento do sistema algumas das alternativas não puderam ser testadas. Os elevados períodos de computação necessários para processar todas as imagens ADNI fizeram com que alguns dos procedimentos não puderam ser aperfeiçoados, pelo que estes podem ser considerados para um trabalho futuro.

Antes de qualquer proposta de melhoria, o problema das máscaras ADNI necessita de ser resolvido. Idealmente, os hipocampos deveriam ser segmentados já no espaço estereotáxico. Como é improvável que o ADNI repita o processo de segmentação, cada uma das transformações aplicadas às máscaras hipocampais deveriam ser verificadas e corrigidas por especialistas, de forma assegurar a correta identificação do hipocampo. O número

elevado de imagens torna esta tarefa difícil, e a alteração das máscaras por especialistas aumentaria a variabilidade entre estudos. Uma outra alternativa para a resolução do problema seria a conversão da resolução das máscaras para a mesma das imagens, antes de aplicar a transformação espacial. Uma forma simples de alcançar esta conversão seria copiar os ficheiros das imagens, e igualar os vóxeis aos valores das coordenadas correspondentes no ficheiro da máscara. O ficheiro resultante teria a resolução da imagem RM e a informação da máscara hipocampal, podendo então ser convertido para o espaço estereotáxico com a transformação do registo linear.

A biblioteca MINC foi usada para todo pré-processamento devido ao facto desta ser única com uma implementação pública do filtro de média não local proposto por Coupé et al. [75]. No entanto, os elevados tempos de computação das suas ferramentas fazem com que o *pipeline* de pré-processamento tenha durações superiores às da extração dos biomarcadores. Os tempos de pré-processamento reportados por Coupé et al. mostram que execução mais rápidas são possíveis [66] pelo que, o uso de outras bibliotecas pode diminuir significativamente o tempo de execução do primeiro passo do método.

O estudo de Yassa e Stark mostra que técnicas de registo não linear, nomeadamente o algoritmo Demons, conseguem melhores resultados no alinhamento de estruturas complexas, como o hipocampo [56]. Assim, e ainda dentro do pré-processamento, métodos de registo não-lineares podem ser explorados de forma a melhorar o registo inicial e, consequentemente, os resultados dos algoritmos SNIPE implementados. O uso de um atlas representativo das populações de interesse também é importante para a obtenção do melhor registo espacial possível [39]. A criação de um atlas a partir da base de dados ADNI pelo método de Fonov et al. exige uma grande quantidade de recursos computacionais, que podem não estar disponíveis. Uma alternativa pode passar pelo uso do atlas para alzheimer disponibilizado pelo Laboratory of Neuro Imaging [104].

Relativamente ao novo método proposto, embora os resultados apresentados não comprovem uma melhoria relativamente ao SNIPE, existem vários caminhos que podem ser seguidos para tentar uma melhor adaptação do algoritmo. Ao contrário do SNIPE, um estudo para determinar os melhores parâmetros do SNIPE-LoG não foi realizado. Por isso, os valores do tamanho dos patches, do tamanho do volume de procura ou do parâmetro de

espalhamento escolhidos podem não ter sido os melhores. A elevada carga computacional do método limita a sua futura utilização em casos reais. Uma forma de acelerar o processamento seria a utilização da informação das bordas como forma de pré-seleção dos *patches*. Para isso, a melhor forma de combinar o resultado do filtro Laplaciano com a medida de similaridade poderia ser estudada. Além disso, a gradação do hipocampo poderia ainda ser limitada a regiões com altos gradientes de intensidades, permitindo simultaneamente reduzir o tempo de computação e resolver o problema das gradações das fronteiras serem atenuadas na média da gradação.

Algumas modificações podem ainda ser sugeridas para melhorar simultaneamente os dois métodos utilizados. Desde de 2012, novas fases de ANDI foram lançadas e novos dados foram disponibilizados. Uma das possibilidades seria o uso das MRI de 3T. É provável que a maior resolução destas imagens permita melhores segmentações e gradações do hipocampo.

É ainda possível estudar novas formas de obter biomarcadores a partir das segmentações e dos mapas de gradação resultantes dos métodos. O valor do volume providencia uma capacidade de discriminação significativa, mas é inadequado para descrever totalmente o efeito da doença na morfologia do hipocampo. Para além do volume, a análise da forma hipocampal também contribuiu para perceber o desenvolvimento da doença [33]. Gerig et al., por exemplo, usaram harmónicos esféricos no hipocampo para detetar mudanças causadas pelo Alzheimer [105].

Na maioria dos processamentos de imagem, é comum o uso de técnicas de pós-processamento. Como nenhum pós-processamento foi proposto por Copé et al., esta é uma área que pode ser desenvolvida. No mapa de gradação, por exemplo, é incoerente que vóxeis vizinhos tenham valores de gradação muito diferentes. Assim, a suavização do mapa de gradação pode ajudar na remoção de vóxeis isolados erradamente classificados.

Finalmente, mudanças no esquema de classificação podem ainda ser consideradas. Vários dos estudos que usam os dados da ADNI para prever a conversão para Alzheimer usam apenas a primeira imagem RM do paciente [106]. Isto torna o grupo pMCI muito heterogéneo no que diz respeito ao tempo de conversão, uma vez que alguns destes sujeitos convertem para Alzheimer ao longo dos 6 meses seguintes enquanto outros têm conversões

diagnosticadas ao fim de 4 anos. Por exemplo, o padrão de atrofia pode ser diferente em pacientes 1 ano antes do diagnóstico comparado com o padrão de 2 anos antes. Eskildsen et al. incorporaram as imagens de 6, 12 e 24 meses antes do diagnóstico AD no seu estudo, e agruparam-nas de acordo com tempo de conversão [106]. Uma metodologia semelhante poderia ser aplicada ao SNIPE e ao SNIPE-LoG.

Bibliografia

- [1] M. Bairu and M. W. Weiner, *Global Clinical Trials for Alzheimer's Disease*. Elsevier, 2014.
- [2] L. C. Lu and J. H. B. M.D., *Alzheimer's Disease*. 2011.
- [3] S. Gauthier, a Leuzy, E. Racine, and P. Rosa-Neto, "Diagnosis and management of Alzheimer's disease: past, present and future ethical issues.," *Prog. Neurobiol.*, vol. 110, pp. 102–13, Nov. 2013.
- [4] S. Saraswathi, B. S. Mahanand, A. Kloczkowski, S. Suresh, and N. Sundararajan, "Detection of onset of Alzheimer ' s Disease from MRI images Using a GA-ELM-PSO Classifier," pp. 42–48, 2013.
- [5] I. Zerr, *UNDERSTANDING ALZHEIMER ' S DISEASE Edited by Inga Zerr. .*
- [6] H. V Bhujle and S. Chaudhuri, "Laplacian based non-local means denoising of MR images with Rician noise.," *Magn. Reson. Imaging*, vol. 31, no. 9, pp. 1599–610, Nov. 2013.
- [7] P. Andersen, R. Morris, D. Amaral, T. Bliss, and J. O'Keefe, *The hippocampus book*. 2006.
- [8] E. Babusikova, A. Evinova, J. Jurecekova, M. Jesenak, and D. Dobrota, "Alzheimer ' s Disease : Definition , Molecular and Genetic Factors," in *Advanced Understanding of Neurodegenerative Diseases*, D. R. C.-C. Chang, Ed. InTech, 2011.
- [9] W. Xu, A. Marseglia, C. Ferrari, and H. Wang, "Alzheimer ' s Disease : A Clinical Perspective," in *Alzheimer's Disease: A Clinical Perspective, Neurodegenerative Diseases*, U. Kishore, Ed. InTech, 2013.
- [10] H. R. Series, *Alzheimer Disease Sourcebook*, 5th ed. Omnigraphics, 2011.
- [11] I. For and I. Of, *Biomarkers for Early Diagnosis of Alzheimers Disease*. New York: Nova Science Publishers, Inc., 2008.
- [12] L. Shargel, S. Wu-Pong, and A. Yu, *Applied Biopharmaceutics and Pharmacokinetics, 5th Edition*, vol. 94. 2004, p. 826.
- [13] MARISA R. BOYD, *Alzheimer's disease: Diagnosis and treatment*. New York: Nova Science Publishers, Inc., 2011.
- [14] C. A. Lemere and E. Masliah, "Can Alzheimer disease be prevented by amyloid-beta immunotherapy?," *Nat. Rev. Neurol.*, vol. 6, pp. 108–119, 2010.
- [15] K. Blennow, H. Zetterberg, J. O. Rinne, S. Salloway, J. Wei, R. Black, M. Grundman, and E. Liu, "Effect of immunotherapy with bapineuzumab on cerebrospinal fluid biomarker levels in patients with mild to moderate Alzheimer disease.," *Arch. Neurol.*, vol. 69, pp. 1002–10, 2012.
- [16] C.-C. Liu, C.-C. Liu, T. Kanekiyo, H. Xu, and G. Bu, "Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy.," *Nat. Rev. Neurol.*, vol. 9, pp. 106–18, 2013.
- [17] Q. Shen, D. a Loewenstein, E. Potter, W. Zhao, J. Appel, M. T. Greig, A. Raj, A. Acevedo, E. Schofield, W. Barker, Y. Wu, H. Potter, and R. Duara, "Volumetric and visual rating of magnetic resonance imaging scans in the diagnosis of amnesic mild cognitive impairment and Alzheimer's disease.," *Alzheimers. Dement.*, vol. 7, no. 4, pp. e101–e108, Jul. 2011.

- [18] E. Board, *Advanced Image Processing Magnetic Resonance Imaging*. CRC Press, 2005.
- [19] P. Prasad, *Magnetic resonance imaging: methods and biologic applications*. New Jersey: Humana Press, 2006, p. 445.
- [20] D. Weishaupt, V. D. Köchli, and B. Marincek, *How does MRI work?*, Second Edi. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003.
- [21] T. P. Naidich, M. Castillo, S. Cha, and J. G. Smirniotopoulos, *Imaging of the Brain: Expert Radiology Series*. Elsevier Health Sciences, 2012, p. 1072.
- [22] Vadim Kuperman, *Magnetic Resonance Imaging Physical Principles and Applications*, Isaak M. Aye. Maryland: Academic Press.
- [23] R. K. Clark, *Anatomy and Physiology: Understanding the Human Body*, 1st ed. Jones & Bartlett Publishers, 2005.
- [24] D. Clark, N. Boutros, and M. Mendez, *The brain and behavior: an introduction to behavioral neuroanatomy*, 3rd ed. 2010.
- [25] E. N. Marieb and K. Hoehn, *Human Anatomy & Physiology*. .
- [26] H. Duvernoy, "The Human Hippocampus. Functional Anatomy, Vascularization and Serial Sections with MRI," *Springer, Berlin*, 2005.
- [27] T. Bartsch, Ed., *The Clinical Neurobiology of the Hippocampus*. Oxford University Press, 2012.
- [28] P. Coupé, J. V. Manjón, V. Fonov, J. Pruessner, M. Robles, and D. L. Collins, "Patch-based segmentation using expert priors: application to hippocampus and ventricle segmentation.," *Neuroimage*, vol. 54, no. 2, pp. 940–54, Jan. 2011.
- [29] R. Wolz, "Automated Extraction of Biomarkers for Alzheimer ' s Disease from Brain Magnetic Resonance Images," 2011.
- [30] J. W. Ashford, A. Rosen, and M. Adamson, "Handbook of Imaging the Alzheimer Brain," *Advances in Alzheimer's Disease*. p. 1. Chady T, Gratkowski S, Takagi T. Electromagneti, 2011.
- [31] K. Partanen, "Volumes of the Entorhinal and Perirhinal Cortices in Alzheimer ' s Disease," *Neurobiol. Aging*, vol. 19, no. 1, pp. 15–22, 1998.
- [32] R. Insausti, K. Juottonen, H. Soininen, A. M. Insausti, K. Partanen, P. Vainio, M. P. Laakso, and A. Pitka, "MR Volumetric Analysis of the Human Entorhinal , Perirhinal , and Temporopolar Cortices," no. April, pp. 659–671, 1998.
- [33] K. Shen, "Automatic Segmentation and Shape Analysis of Human Hippocampus in Alzheimer ' s Disease," UNIVERSITE DE BOURGOGNE, 2011.
- [34] C. Konrad, T. Ukas, C. Nebel, V. Arolt, A. W. Toga, and K. L. Narr, "Defining the human hippocampus in cerebral magnetic resonance images-An overview of current segmentation protocols," *NeuroImage*, vol. 47, pp. 1185–1195, 2009.
- [35] J. Talairach and P. Tournoux, *Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system: an approach to cerebral imaging*, vol. 39. 1988, p. 145.
- [36] A. C. Evans, D. L. Collins, S. R. Mills, E. D. Brown, R. L. Kelly, and T. M. Peters, "3D statistical neuroanatomical models from 305 MRI volumes," *1993 IEEE Conf. Rec. Nucl. Sci. Symp. Med. Imaging Conf.*, 1993.
- [37] D. Collins, "3D Model-based segmentation of individual brain structures from magnetic resonance imaging data," 1994.

- [38] J. Mazziotta, A. Toga, A. Evans, P. Fox, J. Lancaster, K. Zilles, R. Woods, T. Paus, G. Simpson, B. Pike, C. Holmes, L. Collins, P. Thompson, D. MacDonald, M. Iacoboni, T. Schormann, K. Amunts, N. Palomero-Gallagher, S. Geyer, L. Parsons, K. Narr, N. Kabani, G. Le Goualher, D. Boomsma, T. Cannon, R. Kawashima, and B. Mazoyer, "A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM).," *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, vol. 356, pp. 1293–1322, 2001.
- [39] O. T. Carmichael, H. A. Aizenstein, S. W. Davis, J. T. Becker, P. M. Thompson, C. C. Meltzer, and Y. Liu, "Atlas-based hippocampus segmentation in Alzheimer's disease and mild cognitive impairment," *Neuroimage*, vol. 27, pp. 979–990, 2005.
- [40] E. A. Ashton, K. J. Parker, M. J. Berg, and C. W. Chen, "A novel volumetric feature extraction technique with applications to MR images.," *IEEE Trans. Med. Imaging*, vol. 16, pp. 365–371, 1997.
- [41] A. Ghanei, H. Soltanian-Zadeh, and J. P. Windham, "Segmentation of the hippocampus from brain MRI using deformable contours," *Comput. Med. Imaging Graph.*, vol. 22, pp. 203–216, 1998.
- [42] M. Chupin, A. R. Mukuna-Bantumbakulu, D. Hasboun, E. Bardinet, S. Baillet, S. Kinkingnéhun, L. Lemieux, B. Dubois, and L. Garnero, "Anatomically constrained region deformation for the automated segmentation of the hippocampus and the amygdala: Method and validation on controls and patients with Alzheimer's disease," *Neuroimage*, vol. 34, pp. 996–1019, 2007.
- [43] J. Besag, "Digital Image Processing," *J. Appl. Stat.*, vol. 16, pp. 395–407, 1989.
- [44] C. A. Bishop, M. Jenkinson, J. Declerck, and D. Merhof, "Novel Fast Marching for Automated Segmentation of the Hippocampus (FMASH): Method and Validation on Clinical Data," *Neuroimage*, vol. 55, pp. 1009–1019, 2010.
- [45] J. A. Sethian, "A fast marching level set method for monotonically advancing fronts.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 93, pp. 1591–1595, 1996.
- [46] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, 2001.
- [47] D. Shen, S. Moffat, S. M. Resnick, and C. Davatzikos, "Measuring size and shape of the hippocampus in MR images using a deformable shape model.," *Neuroimage*, vol. 15, pp. 422–434, 2002.
- [48] K. O. Babalola, B. Patenaude, P. Aljabar, J. Schnabel, D. Kennedy, W. Crum, S. Smith, T. Cootes, M. Jenkinson, and D. Rueckert, "An evaluation of four automatic methods of segmenting the subcortical structures in the brain," *Neuroimage*, vol. 47, pp. 1435–1447, 2009.
- [49] B. Patenaude, S. M. Smith, D. N. Kennedy, and M. Jenkinson, "A Bayesian model of shape and appearance for subcortical brain segmentation," *Neuroimage*, vol. 56, pp. 907–922, 2011.
- [50] S. Hu and D. L. Collins, "Joint level-set shape modeling and appearance modeling for brain structure segmentation," *Neuroimage*, vol. 36, pp. 672–683, 2007.
- [51] D. L. Collins and A. C. Evans, "Animal: Validation and Applications of Nonlinear Registration-Based Segmentation," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 11, pp. 1271–1294, 1997.

- [52] T. Rohlfing, D. B. Russakoff, and C. R. Maurer, "Performance-based classifier combination in atlas-based image segmentation using expectation-maximization parameter estimation," in *IEEE Transactions on Medical Imaging*, 2004, vol. 23, pp. 983–994.
- [53] R. A. Heckemann, J. V. Hajnal, P. Aljabar, D. Rueckert, and A. Hammers, "Automatic anatomical brain MRI segmentation combining label propagation and decision fusion," *Neuroimage*, vol. 33, pp. 115–126, 2006.
- [54] K. O. Babalola, B. Patenaude, P. Aljabar, J. Schnabel, D. Kennedy, W. Crum, S. Smith, T. F. Cootes, M. Jenkinson, and D. Rueckert, "Comparison and evaluation of segmentation techniques for subcortical structures in brain MRI," *Med. Image Comput. Comput. Assist. Interv.*, vol. 11, pp. 409–416, 2008.
- [55] R. Wolz, P. Aljabar, J. V. Hajnal, A. Hammers, and D. Rueckert, "LEAP: Learning embeddings for atlas propagation," *Neuroimage*, vol. 49, pp. 1316–1325, 2010.
- [56] M. A. Yassa and C. E. L. Stark, "A quantitative evaluation of cross-participant registration techniques for MRI studies of the medial temporal lobe," *Neuroimage*, vol. 44, pp. 319–327, 2009.
- [57] S. K. Warfield, K. H. Zou, and W. M. Wells, "Simultaneous truth and performance level estimation (STAPLE): An algorithm for the validation of image segmentation," *IEEE Trans. Med. Imaging*, vol. 23, pp. 903–921, 2004.
- [58] O. Commowick, S. K. Warfield, and G. Malandain, "Using Frankenstein's creature paradigm to build a patient specific atlas," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2009, vol. 5762 LNCS, pp. 993–1000.
- [59] X. Artaechevarria, A. Muñoz-Barrutia, and C. Ortiz-de-Solórzano, "Combination strategies in multi-atlas image segmentation: Application to brain MR data," *IEEE Trans. Med. Imaging*, vol. 28, pp. 1266–1277, 2009.
- [60] P. Coupé, S. F. Eskildsen, J. V. Manjón, V. S. Fonov, J. C. Pruessner, M. Allard, and D. L. Collins, "NeuroImage : Clinical Scoring by nonlocal image patch estimator for early detection of Alzheimer ' s disease ☆," *Neuroimage*, vol. 1, pp. 141–152, 2012.
- [61] M. R. Sabuncu, B. T. T. Yeo, K. Van Leemput, B. Fischl, and P. Golland, "A generative model for image segmentation based on label fusion," *IEEE Trans. Med. Imaging*, vol. 29, pp. 1714–1729, 2010.
- [62] M. Sdika, "Combining atlas based segmentation and intensity classification with nearest neighbor transform and accuracy weighted vote," *Med. Image Anal.*, vol. 14, pp. 219–226, 2010.
- [63] F. van der Lijn, T. den Heijer, M. M. B. Breteler, and W. J. Niessen, "Hippocampus segmentation in MR images using atlas registration, voxel classification, and graph cuts," *Neuroimage*, vol. 43, pp. 708–720, 2008.
- [64] J. M. Lötjönen, R. Wolz, J. R. Koikkalainen, L. Thurfjell, G. Waldemar, H. Soininen, and D. Rueckert, "Fast and robust multi-atlas segmentation of brain magnetic resonance images," *Neuroimage*, vol. 49, pp. 2352–2365, 2010.
- [65] R. Wolz, P. Aljabar, D. Rueckert, R. A. Heckemann, and A. Hammers, "Segmentation of subcortical structures and the hippocampus in brain MRI using graph-cuts and subject-specific a-priori information," in *Proceedings - 2009 IEEE International*

- Symposium on Biomedical Imaging: From Nano to Macro, ISBI 2009*, 2009, pp. 470–473.
- [66] P. Coupé, S. F. Eskildsen, J. V. Manjón, V. S. Fonov, and D. L. Collins, “Simultaneous segmentation and grading of anatomical structures for patient’s classification: application to Alzheimer’s disease.,” *Neuroimage*, vol. 59, no. 4, pp. 3736–47, Feb. 2012.
- [67] A. Efros and W. Freeman, “Image Quilting for Texture Synthesis and Transfer,” in *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, 2001, pp. 1–6.
- [68] M. Protter, M. Elad, H. Takeda, and P. Milanfar, “Generalizing the nonlocal-means to super-resolution reconstruction,” *IEEE Trans. Image Process.*, vol. 18, pp. 36–51, 2009.
- [69] A. Criminisi, P. Pérez, and K. Toyama, “Region filling and object removal by exemplar-based image inpainting,” *IEEE Trans. Image Process.*, vol. 13, pp. 1200–1212, 2004.
- [70] a. Buades and B. Coll, “A non-local algorithm for image denoising,” *Comput. Vis. Pattern*, vol. 2, pp. 60–65, 2005.
- [71] A. Buades, B. Coll, and J. M. Morel, “Image Denoising Methods. A New Nonlocal Principle,” *SIAM Review*, vol. 52. pp. 113–147, 2010.
- [72] P. Aljabar, R. A. Heckemann, A. Hammers, J. V. Hajnal, and D. Rueckert, “Multi-atlas based segmentation of brain images: Atlas selection and its effect on accuracy,” *Neuroimage*, vol. 46, pp. 726–738, 2009.
- [73] D. L. Collins and J. C. Pruessner, “Towards accurate, automatic segmentation of the hippocampus and amygdala from MRI by augmenting ANIMAL with a template library and label fusion,” *Neuroimage*, vol. 52, pp. 1355–1366, 2010.
- [74] T. Rohlfing, R. Brandt, R. Menzel, and C. R. Maurer, “Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains,” *Neuroimage*, vol. 21, pp. 1428–1442, 2004.
- [75] P. Coupe, P. Yger, S. Prima, P. Hellier, C. Kervrann, and C. Barillot, “An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images.,” *IEEE Trans. Med. Imaging*, vol. 27, no. 4, pp. 425–41, Apr. 2008.
- [76] N. Wiest-Daesslé, S. Prima, P. Coupé, S. P. Morrissey, and C. Barillot, “Rician noise removal by non-local means filtering for low signal-to-noise ratio MRI: Applications to DT-MRI,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2008, vol. 5242 LNCS, pp. 171–179.
- [77] J. G. Sled, a P. Zijdenbos, and a C. Evans, “A nonparametric method for automatic correction of intensity nonuniformity in MRI data.,” *IEEE Trans. Med. Imaging*, vol. 17, no. 1, pp. 87–97, Feb. 1998.
- [78] R. G. Boyes, J. L. Gunter, C. Frost, A. L. Janke, T. Yeatman, D. L. G. Hill, M. A. Bernstein, P. M. Thompson, M. W. Weiner, N. Schuff, G. E. Alexander, R. J. Killiany, C. DeCarli, C. R. Jack, and N. C. Fox, “Intensity non-uniformity correction using N3 on 3-T scanners with multichannel phased array coils,” *Neuroimage*, vol. 39, pp. 1752–1762, 2008.
- [79] J. B. Arnold, J. S. Liow, K. A. Schaper, J. J. Stern, J. G. Sled, D. W. Shattuck, A. J. Worth, M. S. Cohen, R. M. Leahy, J. C. Mazziotta, and D. A. Rottenberg, “Qualitative and quantitative evaluation of six algorithms for correcting intensity nonuniformity effects.,” *Neuroimage*, vol. 13, pp. 931–943, 2001.

- [80] W. Zheng, M. W. L. Chee, and V. Zagorodnov, "Improvement of brain segmentation accuracy by optimizing non-uniformity correction using N3," *Neuroimage*, vol. 48, pp. 73–83, 2009.
- [81] M. Jenkinson and S. Smith, "A global optimisation method for robust affine registration of brain images," *Med. Image Anal.*, vol. 5, pp. 143–156, 2001.
- [82] "Jim Version 6.0 Manual." [Online]. Available: <http://www.xinapse.com/Manual/xyshear.gif>. [Accessed: 22-Oct-2014].
- [83] L. G. Nyú and J. K. Udupa, "On standardizing the MR image intensity scale," *Magn. Reson. Med.*, vol. 42, pp. 1072–1081, 1999.
- [84] M. Chupin, E. Gérardin, R. Cuingnet, C. Boutet, L. Lemieux, S. Lehericy, H. Benali, L. Garnero, and O. Colliot, "Fully automatic hippocampus segmentation and classification in Alzheimer's disease and mild cognitive impairment applied on data from ADNI," *Hippocampus*, vol. 19, pp. 579–587, 2009.
- [85] R. Wolz, V. Julkunen, J. Koikkalainen, E. Niskanen, D. P. Zhang, D. Rueckert, H. Soininen, and J. Lötjönen, "Multi-method analysis of MRI images in early diagnostics of Alzheimer's disease.," *PLoS One*, vol. 6, no. 10, p. e25446, Jan. 2011.
- [86] V. Fonov, A. C. Evans, K. Botteron, C. R. Almli, R. C. McKinstry, and D. L. Collins, "Unbiased average age-appropriate atlases for pediatric studies.," *Neuroimage*, vol. 54, no. 1, pp. 313–27, Jan. 2011.
- [87] J. Ashburner and K. Friston, "Multimodal image coregistration and partitioning—a unified framework.," *Neuroimage*, vol. 6, pp. 209–217, 1997.
- [88] S. M. Smith, M. Jenkinson, M. W. Woolrich, C. F. Beckmann, T. E. J. Behrens, H. Johansen-Berg, P. R. Bannister, M. De Luca, I. Drobnjak, D. E. Flitney, R. K. Niazy, J. Saunders, J. Vickers, Y. Zhang, N. De Stefano, J. M. Brady, and P. M. Matthews, "Advances in functional and structural MR image analysis and implementation as FSL," in *NeuroImage*, 2004, vol. 23.
- [89] S. F. Eskildsen, P. Coupé, V. Fonov, J. V. Manjón, K. K. Leung, N. Guizard, S. N. Wassef, L. R. Østergaard, and D. L. Collins, "BEaST: Brain extraction based on nonlocal segmentation technique," *Neuroimage*, vol. 59, pp. 2362–2373, 2012.
- [90] J. Pruessner, S. Köhler, J. Crane, and C. Lord, "Entorhinal and parahippocampal cortex from high-resolution MR images?: Considering the variability of the collateral sulcus," *Cereb. Cortex*, pp. 1342–1353, 2002.
- [91] W. Contributors, "MINC," *Wikibooks, The Free Textbook Project*. [Online]. Available: <http://en.wikibooks.org/w/index.php?title=MINC&oldid=2206989>. [Accessed: 17-Nov-2014].
- [92] W. Schroeder, "The ITK Software Guide Second Edition Updated for ITK version 2 . 4," *FEBS Lett.*, vol. 525, pp. 53–58, 2005.
- [93] S. G. Mueller, M. W. Weiner, L. J. Thal, R. C. Petersen, C. Jack, W. Jagust, J. Q. Trojanowski, A. W. Toga, and L. Beckett, "The Alzheimer's disease neuroimaging initiative," *Neuroimaging Clinics of North America*, vol. 15. pp. 869–877, 2005.
- [94] P. S. Aisen, R. C. Petersen, M. C. Donohue, A. Gamst, R. Raman, R. G. Thomas, S. Walter, J. Q. Trojanowski, L. M. Shaw, L. A. Beckett, C. R. Jack, W. Jagust, A. W. Toga, A. J. Saykin, J. C. Morris, R. C. Green, and M. W. Weiner, "Clinical core of the Alzheimer's disease neuroimaging initiative: Progress and plans," *Alzheimer's and Dementia*, vol. 6. pp. 239–246, 2010.

- [95] R. C. Petersen, P. S. Aisen, L. A. Beckett, M. C. Donohue, A. C. Gamst, D. J. Harvey, C. R. Jack, W. J. Jagust, L. M. Shaw, A. W. Toga, J. Q. Trojanowski, and M. W. Weiner, "Alzheimer's Disease Neuroimaging Initiative (ADNI): Clinical characterization," *Neurology*, vol. 74, pp. 201–209, 2010.
- [96] B. T. Wyman, D. J. Harvey, K. Crawford, M. a Bernstein, O. Carmichael, P. E. Cole, P. K. Crane, C. DeCarli, N. C. Fox, J. L. Gunter, D. Hill, R. J. Killiany, C. Pachai, A. J. Schwarz, N. Schuff, M. L. Senjem, J. Suhy, P. M. Thompson, M. Weiner, and C. R. Jack, "Standardization of analysis sets for reporting results from ADNI MRI data.," *Alzheimers. Dement.*, vol. 9, no. 3, pp. 332–7, May 2013.
- [97] S. Aruoba and J. Fernández-Villaverde, "A Comparison of Programming Languages in Economics," *NBER Work. Pap. Ser.*, vol. 3, pp. 0–16, 2014.
- [98] R. Sebesta, *Concepts of programming languages*. The Press Syndicate of the University of Cambridge, 2002.
- [99] M. Boccardi, M. Bocchetta, L. G. Apostolova, J. Barnes, G. Bartzokis, G. Corbetta, C. DeCarli, L. deToledo-Morrell, M. Firbank, R. Ganzola, L. Gerritsen, W. Henneman, R. J. Killiany, N. Malykhin, P. Pasqualetti, J. C. Pruessner, A. Redolfi, N. Robitaille, H. Soininen, D. Tolomeo, L. Wang, C. Watson, H. Wolf, H. Duvernoy, S. Duchesne, C. R. Jack, and G. B. Frisoni, "Delphi definition of the EADC-ADNI Harmonized Protocol for hippocampal segmentation on magnetic resonance," *Alzheimer's and Dementia*, 2014.
- [100] H. V. Bhujle and S. Chaudhuri, "Laplacian based non-local means denoising of MR images with Rician noise," *Magn. Reson. Imaging*, vol. 31, pp. 1599–1610, 2013.
- [101] A. Akselrod-Ballin, M. Galun, M. J. Gomori, R. Basri, and A. Brandt, "Atlas guided identification of brain structures by combining 3D segmentation and SVM classification.," *Med. Image Comput. Comput. Assist. Interv.*, vol. 9, pp. 209–216, 2006.
- [102] K. J. van Stralen, V. S. Stel, J. B. Reitsma, F. W. Dekker, C. Zoccali, and K. J. Jager, "Diagnostic methods I: sensitivity, specificity, and other measures of accuracy.," *Kidney Int.*, vol. 75, pp. 1257–1263, 2009.
- [103] C. Carbno, "Business Statistics: Contemporary Decision Making," *Technometrics*, vol. 49, pp. 495–496, 2007.
- [104] "LONI Atlases." [Online]. Available: <http://www.loni.usc.edu/atlasses>.
- [105] G. Gerig, M. Styner, D. Jones, D. Weinberger, and J. Lieberman, "Shape analysis of brain ventricles using SPHARM," *Proc. IEEE Work. Math. Methods Biomed. Image Anal. (MMBIA 2001)*, 2001.
- [106] S. F. Eskildsen, P. Coupé, D. García-Lorenzo, V. Fonov, J. C. Pruessner, and D. L. Collins, "Prediction of Alzheimer's disease in subjects with mild cognitive impairment from the ADNI cohort using patterns of cortical thinning," *Neuroimage*, vol. 65, pp. 511–521, 2013.

Anexo I – Falha do pré-processamento

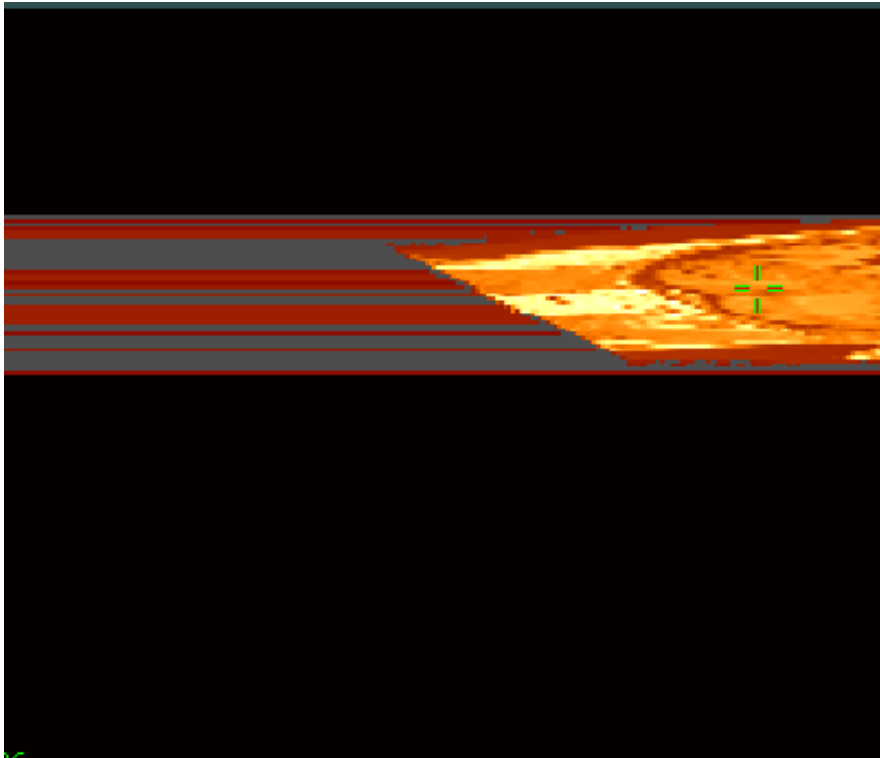


Figura 1- Vista axial da imagem resultante do pré-processamento do sujeito 1244.

Anexo II – Resultados dos testes estatísticos

Tabela 1- Valores-p do teste *Wilcoxon-Mann-Whitney* resultantes das comparações do volume e da gradação média, obtidos com a implementação do SNIPE, das diferentes populações

Comparação	Volume	Gradação
CN vs. sMCI	2,86E-12	4,14E-08
CN vs. pMCI	2,20E-16	2,20E-16
CN vs. AD	2,20E-16	2,20E-16
sMCI vs. pMCI	8,24E-06	7,37E-22
sMCI vs. AD	8,31E-13	2,20E-16
pMCI vs. AD	8,62E-03	3,11E-02

Tabela 2- Valores-p do teste *Wilcoxon-Mann-Whitney* resultantes das comparações do volume e da gradação média, obtidos com a implementação do SNIPE, das populações sMCI e pMCI construídas em 2012 e em 2014

População	Volume	Gradação
sMCI	0,9556	0,832
pMCI	0,5104	0,239

Tabela 3- Valores-p do teste *Wilcoxon-Mann-Whitney* resultantes das comparações do volume e da gradação média, obtidos com a implementação do SNIPE-LoG, das diferentes populações

Comparação	Volume	Gradação
CN vs. sMCI	2,86E-12	5,17E-06
CN vs. pMCI	2,20E-16	2,20E-16
CN vs. AD	2,20E-16	2,20E-16
sMCI vs. pMCI	8,24E-16	3,64E-11
sMCI vs. AD	8,31E-13	2,20E-16
pMCI vs. AD	8,62E-03	3,02E-02

Tabela 4- Valores-p do teste *Wilcoxon-Mann-Whitney* resultantes das comparações do volume e da gradação média, obtidos com a implementação do SNIPE-LoG, das populações sMCI e pMCI construídas em 2012 e em 2014

População	Volume	Gradação
sMCI	0,4915	0,1795
pMCI	0,3262	0,3106