

UNIVERSIDADE FEDERAL DE SÃO CARLOS– UFSCAR
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA– CCET
DEPARTAMENTO DE COMPUTAÇÃO– DC
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO– PPGCC

Marco Antônio Bernardi Grivol

**Assimetrias Inter-hemisféricas no
Hipocampo e sua Associação com a
Progressão da Doença de Alzheimer**

São Carlos - SP
Agosto de 2025

Marco Antônio Bernardi Grivol

**Assimetrias Inter-hemisféricas no
Hipocampo e sua Associação com a
Progressão da Doença de Alzheimer**

Dissertação apresentada ao Programa de Pós-Graduação em
Ciência da Computação do Centro de Ciências Exatas e de
Tecnologia da Universidade Federal de São Carlos, como parte
dos requisitos para a obtenção do título de Mestre em Ciência
da Computação.

Área de concentração: Metodologias e Técnicas de Computação

Orientador: Prof. Dr. Ricardo José Ferrari

São Carlos - SP

Agosto de 2025



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Exatas e de Tecnologia
Programa de Pós-Graduação em Ciência da Computação

Folha de Aprovação

Defesa de Dissertação de Mestrado do candidato Marco Antonio Bernardi Grivol, realizada em 29/07/2025.

Comissão Julgadora:

Prof. Dr. Ricardo José Ferrari (UFSCar)

Prof. Dr. Jurandy Gomes de Almeida Junior (UFSCar)

Prof. Dr. Leandro Alves Neves (UNESP)

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa de Pós-Graduação em Ciência da Computação.

Agradecimentos

Agradeço aos meus pais, Otávio e Cristiane, pelo apoio em minhas escolhas, por estarem sempre presentes oferecendo todo o seu amor e dedicação.

Agradeço ao meu orientador, Prof. Dr. Ricardo José Ferrari, por compartilhar seus conhecimentos e conselhos. Sua presença constante, disposição para ajudar e as discussões construtivas realizadas com muita paciência, dedicação e sabedoria.

Também expresso minha gratidão a todos os membros do BIP-Group que compartilharam experiências e conteúdos de forma aberta e colaborativa.

Resumo

A atrofia cerebral é um processo progressivo típico do envelhecimento natural. Em doenças neurodegenerativas, como a doença de Alzheimer (DA), esse processo é acelerado, manifestando-se em sintomas relacionados à perda de memória, dificuldades de raciocínio e mudanças no comportamento. Atualmente, não existe cura para a DA, sendo o tratamento precoce a forma mais eficaz de minimizar ou retardar os sintomas. Além dos testes clínicos e do histórico familiar, as imagens de Ressonância Magnética (RM) do cérebro são fundamentais no diagnóstico, pois permitem a avaliação visual da atrofia cerebral, que pode ser comparada com a atrofia associada ao envelhecimento natural, e ajudam a excluir outras possíveis comorbidades de demências tratáveis. Adicionalmente, as alterações inter-hemisféricas estão associadas à DA e podem ser observadas em imagens de RM. Estudos recentes têm avaliado novas formas de quantificar essa assimetria com o objetivo de desenvolver biomarcadores para o auxílio-diagnóstico da DA, e treinar classificadores para o diagnóstico automático. No entanto, mesmo com bons resultados, a inspeção visual das imagens ainda é a forma mais comum de avaliar a presença da DA em imagens de RM. Nesse sentido, este trabalho desenvolve uma nova metodologia para quantificar as assimetrias inter-hemisféricas do hipocampo com base em descritores geométricos e texturais, com foco na predição da progressão do Comprometimento Cognitivo Leve (CCL) estável para o CCL progressivo. A abordagem utiliza a técnica *shape context* para extrair histogramas log-esféricos do mapa de energia acumulada da Transformada *Wavelet* Complexa em Árvore-Dupla (DT-CWT) 3D na região do hipocampo. Esses descritores, em conjunto com atributos volumétricos, são combinados em uma Rede Neural Convolucional Siamesa projetada para modelar diferenças inter-hemisféricas. A metodologia proposta obteve F1-score de $0,845 \pm 0,01$ na distinção entre casos CCL estáveis vs. progressivos. Experimentos adicionais de classificação binária envolvendo indivíduos cognitivamente normais (CN), CCL e DA também foram conduzidos, com resultados consistentes: CN vs. DA (F1-score = $0,872 \pm 0,05$), CN vs. CCL ($0,757 \pm 0,03$) e CCL vs. DA ($0,731 \pm 0,05$). Dessa forma, esses resultados demonstram a influência de alterações

inter-hemisféricas na região do hipocampo durante a progressão da DA e a eficácia desses atributos na predição e prognóstico da doença.

Palavras-chave: Detecção de Pontos Salientes, Transformada *Wavelet* Complexa de Árvore Dupla, Doença de Alzheimer, Assimetrias Inter-Hemisféricas no Encéfalo, Comprometimento Cognitivo Leve.

Abstract

Brain atrophy is a progressive process typically associated with natural aging. In neurodegenerative diseases such as Alzheimer’s disease (AD), this process is accelerated, manifesting in symptoms related to memory loss, reasoning difficulties, and behavioral changes. Currently, there is no cure for AD, and early treatment remains the most effective way to mitigate or delay symptoms. In addition to clinical tests and family history, structural Magnetic Resonance Imaging (MRI) of the brain plays a fundamental role in diagnosis, as it enables visual assessment of brain atrophy—allowing comparison with age-related atrophy—and assists in ruling out other possible comorbidities from treatable dementias. Moreover, inter-hemispheric alterations are associated with AD and can be observed in MRI scans. Recent studies have investigated new methods for quantifying such asymmetries in order to develop biomarkers for supporting AD diagnosis and to train classifiers for automated diagnosis. However, despite promising results, visual inspection remains the most common approach for evaluating AD presence in MRI images. In this context, this work proposes a new methodology to quantify inter-hemispheric hippocampal asymmetries based on geometric and textural descriptors, focusing on predicting the progression from stable to progressive Mild Cognitive Impairment (MCI). The approach employs the shape context technique to extract log-spherical histograms from the accumulated energy map of the 3D Dual-Tree Complex Wavelet Transform (DT-CWT) in the hippocampal region. These descriptors, combined with volumetric features, are fed into a Convolutional Siamese neural network specifically designed to model inter-hemispheric differences. The proposed methodology achieved an F1-score of 0.845 ± 0.01 in distinguishing stable from progressive MCI cases. Additional binary classification experiments involving cognitively normal (CN), MCI, and AD subjects were also conducted, with consistent results: CN vs. AD (F1-score = 0.872 ± 0.05), CN vs. MCI (0.757 ± 0.03), and MCI vs. AD (0.731 ± 0.05). These findings demonstrate the influence of inter-hemispheric hippocampal alterations during AD progression and the effectiveness of these features for disease prediction and prognosis.

Keywords: Keypoint Detection, Dual-Tree Complex Wavelet Transform, Alzheimer's Disease, Inter-Hemispheric Brain Asymmetries, Mild Cognitive Impairment.

Lista de ilustrações

Figura 1 – Contínuo da DA.	30
Figura 2 – Imagem de RM ponderada em T2, mostrando na imagem ampliada os diferentes tecidos envolvidos nos tempos T2.	34
Figura 3 – Representação das <i>wavelets</i> no domínio do espaço (primeira linha) e o suporte das wavelets do espectro de Fourier.	42
Figura 4 – Exemplo de segmentação de imagem de RM do encéfalo em CSF, GM, WM e <i>background</i>	43
Figura 5 – Histograma log-esférico utilizado para calcular o descritor <i>shape context</i>	44
Figura 6 – Índice de assimetria em diferentes estruturas anatômicas.	59
Figura 7 – Diagrama de rosas da diferença absoluta entre os vetores de características do hemisfério esquerdo e direito.	62
Figura 8 – Diagrama da etapa de normalização, redimensionamento e detecção dos pontos salientes localizados no sulco limitado por dois pontos.	74
Figura 9 – Fluxograma das principais etapas deste projeto de pesquisa.	86
Figura 10 – Informações demográficas das imagens da base ADNI.	87
Figura 11 – Informações demográfica das imagens de base IXI.	88
Figura 12 – Principais etapas empregadas na quantificação dos tecidos cerebrais	91
Figura 13 – Imagem sintética ilustrativa do detector DT-CWT. Esquerda: imagem original. Direita: pontos detectados e sobrepostos na imagem original.	92
Figura 14 – Mapa de energia acumulada A para sinais 2D.	92
Figura 15 – Fluxograma das principais técnicas necessárias para construção do mapa de energia acumulada.	93
Figura 16 – Fluxograma da sequência de etapas utilizadas para a detecção dos pontos estáveis na população CN.	95
Figura 17 – Principais etapas empregadas na extração dos descritores <i>shape context</i>	96
Figura 18 – Arquitetura SCNN multimodal.	99

Figura 19 – Visualização comparativa entre a imagem original (primeira linha) e a imagem preprocessada (segunda linha) nas vistas axial, coronal e sagital.	105
Figura 20 – Resultado do corregristo afim entre a imagem de referência SPL/NAC e uma imagem da base ADNI.	106
Figura 21 – Resultado da extração do encéfalo com a técnica SynthStrip.	107
Figura 22 – Visualização comparativa para a redução de ruído com a técnica NLM em três vistas.	108
Figura 23 – Resultado da correção do efeito de <i>bias field</i> com a técnica N4.	109
Figura 24 – Exemplo comparativo de histogramas de intensidade antes (primeira linha) e depois (segunda linha) da normalização FCM, com $c_1 = 1000$	109
Figura 25 – Pontos detectados em três imagens sintéticas geradas por um modelo paramétrico.	111
Figura 26 – Projeções axial, coronal e sagital do atlas probabilístico criado para as regiões hipocampais, a partir dos pontos salientes detectados na população CN.	112
Figura 27 – Distribuições de densidade da métrica de assimetria inter-hemisférica para os grupos CN e DA, para cada uma das quatro estratégias de posicionamento do descritor.	114
Figura 28 – Curvas ROC e seus respectivos valores de AUC para a tarefa de classificação CN vs. CCL.	117
Figura 29 – Curvas ROC para a classificação CN vs. DA.	119
Figura 30 – Curvas ROC para a classificação CCL vs. DA.	120
Figura 31 – Curvas ROC para a classificação prognóstica CCLe vs. CCLp.	122

Lista de tabelas

Tabela 1 – Tabela de Contingência para o Teste de McNemar	52
Tabela 2 – Comparativo dos Trabalhos Correlatos em Análise de Assimetria Cerebral.	68
Tabela 3 – Valores médios do coeficiente de Dice em 129 imagens da base ADNI. .	71
Tabela 4 – Valores de FP e FN sob a influência de ruídos.	72
Tabela 5 – Comparativo dos Trabalhos Correlatos em Detecção de Pontos Salientes 3D.	77
Tabela 6 – Comparativo dos Trabalhos Correlatos em Predição da Progressão do CCL para a DA.	82
Tabela 7 – Informações demográficas das imagens de RM da base ADNI.	87
Tabela 8 – Arquitetura detalhada da DT-CWT SCNN. A tabela descreve os blocos convolucionais do módulo codificador e as camadas do módulo classificador.	100
Tabela 9 – Valores médios da distância (d) entre o ponto ideal e o ponto detectado e quantidade de FPs e FNs que ocorreram nos conjuntos, avaliados em crescentes níveis de degradação com ruído (σ_r) e <i>bias field</i> (BF). . . .	111
Tabela 10 – Análise de Separação de Classe para as Estratégias de Posicionamento do Descritor.	113
Tabela 11 – Resultados da classificação entre indivíduos CN e com CCL.	116
Tabela 12 – Resultados da classificação entre indivíduos CN e com DA.	118
Tabela 13 – Resultados da classificação CCL vs. DA.	120
Tabela 14 – Resultados da classificação prognóstica entre CCLe e CCLp.	122
Tabela 15 – Resultados do Teste de McNemar para a Classificação CN vs. CCL . .	123
Tabela 16 – Resultados do Teste de McNemar para a Classificação CN vs. DA . .	124
Tabela 17 – Resultados do Teste de McNemar para a Classificação CCL vs. DA . .	124
Tabela 18 – Resultados do Teste de McNemar para a Classificação CCLe vs. CCLp	124

Lista de siglas

ADNI Alzheimer's Disease Neuroimaging Initiative

ADRDA Alzheimer's Disease and Related Disorders Association

ANTs Advanced Normalization Tools

BF Bias Field

CAAT Class Activation Attention Transfer

CAM Class Activation Maps

CCL Comprometimento Cognitivo Leve

CCLe Comprometimento Cognitivo Leve estável

CCLp Comprometimento Cognitivo Leve progressivo

CN Cognitivamente Normal

CSF Cerebral Spinal Fluid

CWT Transformada Wavelet Complexa

CCA Córtex Cingulado Anterior

DA Doença de Alzheimer

DNN Deep Neural Network

DT-CWT Dual-Tree Complex Wavelet Transform

DWT Transformada Wavelet Discreta

EM Expectation-Maximization

ENIGMA Enhancing NeuroImaging Genetics through Meta-Analysis

FCM Fuzzy C-Means

GM Grey Matter

MRF Markov Random Field

MSP Mid-Sagittal Plane

MTA Medial Temporal lobe Atrophy

MLP Multilayer Perceptron

N3 Nonparametric Nonuniform Intensity Normalization

NINCDS National Institute of Neurological and Communicative Disorders and Stroke

NLM NonLocal Means

OASIS Open Access Series of Imaging Studies

PC Phase Congruency

PET Positron Emission Tomography

RF Radiofrequência

RM Ressonância Magnética

RNA Rede Neural Artificial

SC Shape Context

SNR Signal-to-Noise Ratio

VLAPP Variante Logopênica da Afasia Progressiva Primária

WLS Weighted Least-Squares

WM White Matter

Sumário

1	INTRODUÇÃO	23
1.1	Contextualização e motivação	23
1.2	Objetivos	26
1.2.1	Objetivo geral	26
1.2.2	Objetivos específicos	26
1.3	Organização do trabalho	27
2	FUNDAMENTAÇÃO TEÓRICA	29
2.1	Doença de Alzheimer	29
2.1.1	Causas e fisiopatologia	30
2.1.2	Sintomas e diagnóstico	30
2.1.3	Tratamentos	31
2.2	Assimetrias inter-hemisféricas	32
2.2.1	Assimetria na apresentação e progressão da DA	32
2.3	Imagens de ressonância magnética	32
2.3.1	Princípios de formação da imagem	33
2.3.2	Contraste da imagem	33
2.4	Deteção do Plano Mediano	34
2.5	Pré-processamento	35
2.5.1	Redução de ruído	36
2.5.2	Correção do efeito <i>bias field</i>	36
2.5.3	Normalização de intensidade	36
2.5.4	Alinhamento espacial	37
2.5.5	Extração do encéfalo	38
2.6	Transformadas <i>Wavelets</i>	39
2.6.1	Transformada <i>wavelet</i> discreta (DWT)	39
2.6.2	Transformada <i>wavelet</i> complexa (CWT)	40

2.6.3	Transformada <i>wavelet</i> complexa de árvore dupla bidimensional (DT-CWT 2D)	40
2.6.4	Transformada <i>wavelet</i> complexa de árvore dupla tridimensional (DT-CWT 3D)	41
2.7	Segmentação de tecidos do encéfalo	42
2.8	Shape context	44
2.8.1	Correspondências entre descritores	45
2.9	XGBoost	46
2.9.1	Gradient Tree Boosting	46
2.9.2	O Algoritmo XGBoost e sua Função de Objetivo Regularizada	47
2.9.3	Otimizações e Escalabilidade	47
2.10	Redes Neurais	48
2.10.1	Arquitetura e Componentes Fundamentais	48
2.10.2	Aprendizado e Otimização	49
2.10.3	Redes Neurais Siamesas	50
2.11	Capacidade, <i>Overfitting</i> e Regularização	51
2.12	Análise estatística	51
2.12.1	Teste de McNemar	52
3	REVISÃO BIBLIOGRÁFICA	55
3.1	Assimetrias inter-hemisféricas	55
3.1.1	Trabalhos correlatos	56
3.1.2	Discussão	66
3.2	Pontos salientes 3D	68
3.2.1	Trabalhos correlatos	68
3.2.2	Discussão	75
3.3	Progressão do CCL para a DA	77
3.3.1	Trabalhos correlatos	78
3.3.2	Discussão	81
3.4	Considerações finais	83
4	MATERIAIS E MÉTODOS	85
4.1	Enquadramento	85
4.2	Base de dados	86
4.2.1	ADNI	86
4.2.2	IXI	87
4.2.3	SPL/NAC	88
4.3	Pré-processamento das imagens	88
4.3.1	Alinhamento espacial	89
4.3.2	Extração do encéfalo (<i>Skull Stripping</i>)	89

4.3.3	Redução de ruído (<i>denoising</i>)	89
4.3.4	Correção do <i>bias field</i>	90
4.3.5	Normalização de intensidade	90
4.3.6	Identificação da região de interesse anatômica	90
4.4	Segmentação e quantificação dos tecidos cerebrais	90
4.5	Detecção de pontos salientes com a DT-CWT	91
4.6	Atlas probabilístico de pontos salientes	94
4.7	Descritor <i>shape context</i>	96
4.7.1	Posicionamento do descritor <i>shape context</i>	96
4.8	Classificador XGBoost	98
4.9	Modelagem da neural Siamesa	98
4.9.1	Modelo de Referência	100
4.10	Atributos inter-hemisféricos	101
4.11	Aplicação do teste de McNemar	102
4.12	Considerações finais	103
5	RESULTADOS E DISCUSSÕES	105
5.1	Pré-processamento	105
5.2	Avaliação do detector de pontos salientes em imagens sintéticas	
	3D	110
5.3	Detecção de pontos estáveis na população CN	111
5.4	Posicionamento do Atlas nas Imagens da Base ADNI	112
5.5	Definição dos Parâmetros do <i>Shape Context</i>	112
5.6	Resultados da classificação	114
5.6.1	Configuração Experimental	114
5.7	Classificação CN vs. CCL	115
5.8	Classificação CN vs. DA	117
5.9	Classificação CCL vs. DA	119
5.10	Classificação CCLe vs. CCLp	120
5.11	Análise de Significância Estatística	123
5.12	Considerações Finais	124
6	CONCLUSÕES	127
6.1	Trabalhos Futuros	128
	REFERÊNCIAS	131

Capítulo 1

Introdução

1.1 Contextualização e motivação

A Doença de Alzheimer (DA) é a forma mais prevalente de demência, constituindo aproximadamente de 60 a 80% de todos os casos (Alzheimer's Association, 2023). A demência, um termo amplo que descreve uma variedade de condições e doenças resultantes da morte ou disfunção neuronal, afeta a memória, o comportamento e a capacidade de raciocínio claro. Indivíduos com DA sofrem um declínio gradual na capacidade de reter novas informações, devido ao dano nas regiões cerebrais responsáveis pela formação de novas memórias. À medida que a doença avança, a capacidade cognitiva diminui progressivamente, resultando em dificuldades nas atividades diárias, como caminhar e engolir (Alzheimer's Association, 2023).

Primeiramente observada por Alois Alzheimer no ano de 1911, dois tipos de mudanças no cérebro são características fundamentais que explicam o funcionamento neuropatológico da doença: (a) lesões positivas ocasionadas do acúmulo de placas amiloides e neurofibrilares; (b) lesões negativas caracterizadas pela neurodegeneração advinda da morte neuronal e sináptica (HOU et al., 2019; BREIJYEH; KARAMAN, 2020). Embora o envelhecimento seja considerado o maior fator de risco na DA (World Health Organization, 2019; HOU et al., 2019), estudos indicam que a doença não é uma consequência inevitável do envelhecimento e, em parte da população talvez seja possível prevenir ou reduzir seu desenvolvimento ao considerar fatores de risco como hipertensão, obesidade, diabetes e tabagismo (RASMUSSEN; LANGERMAN, 2019). Ademais, a formação de placas senis pode ocorrer até trinta anos antes da manifestação dos sintomas clínicos (OOSTVEEN; LANGE, 2021). Portanto, dada a natureza progressiva, e muitas vezes insidiosa da DA, o diagnóstico precoce e preciso é fundamental para a gestão eficaz da doença, na medida

em que o tratamento visa minimizar e retardar os sintomas (PRINCE et al., 2015; HOU et al., 2019; Alzheimer's Association, 2023).

Estabelecido em 1984, o National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) e Alzheimer's Disease and Related Disorders Association (ADRDA) definiram os critérios para o diagnóstico clínico da DA (MCKHANN et al., 1984). Com o avanço nas tecnologias de imageamento médico, em 2010, os critérios NINCDS-ADRDA foram revisados para a inclusão de biomarcadores que definem os estágios da doença e formalizam os critérios para a categorização da doença em três estágios distintos (JACK et al., 2011). Dentre os biomarcadores definidos, a DA pode ser diagnosticada em imagens de Ressonância Magnética (RM) observando-se padrões específicos de atrofia decorrentes da perda de tecido e morte neuronal (JACK et al., 2011; KNOPMAN et al., 2021). Assim, além de auxiliar no diagnóstico médico, biomarcadores presentes em imagens de RM podem excluir outras possíveis comorbidades de demências tratáveis ou indicar a presença de comorbidades que agravam os sintomas, como a doença cerebrovascular (GAUTHIER et al., 2021a; TAKIZAWA et al., 2015). Um exemplo de biomarcador é a escala Medial Temporal lobe Atrophy (MTA), ou escala de Schelten, (SCHELTENS et al., 1992), que é utilizada por médicos para avaliar visualmente a atrofia na região hipocampal como um indicativo da presença e progressão da DA (SCHELTENS et al., 1992; GAUTHIER et al., 2021a).

A assimetria inter-hemisférica é uma característica fundamental para a organização cerebral humana, com implicações em diversas funções cognitivas e perceptivas, incluindo linguagem, processamento facial e raciocínio visoespacial (KONG et al., 2018). Mudanças na assimetria estão fortemente associadas com o envelhecimento natural e podem apresentar alterações devido a fatores como sexo e desenvolvimento motor da mão dominante (KONG et al., 2018). Um exemplo de mudança na assimetria decorrente do desenvolvimento motor pode ser observada em violinistas, cujas mãos esquerdas dedilham continuamente as cordas enquanto tocam, possuem áreas expandidas no hemisfério direito responsáveis pela coordenação dessa atividade. Quanto maior o tempo dedicado à prática, mais significativo é o aumento dessa área (ELBERT et al., 1995). Com isso, um número crescente de estudos recentes buscam mapear e quantificar as assimetrias médias encontradas em indivíduos Cognitivamente Normal (CN) (KUO; MASSOUD, 2022).

Estudos recentes com imagens de RM têm explorado alterações na assimetria estrutural hemisférica na presença de transtornos psiquiátricos (KONG et al., 2022) e condições patológicas (NEMATİ et al., 2023). Na DA, as alterações na assimetria podem ser caracterizadas por sintomas clínicos que se manifestam inicialmente como dificuldades de memória verbais ou deficit nas funções visoespaciais correlacionados com o hemisfério mais afetado pela neurodegeneração, sendo o hemisfério esquerdo aquele que comumente apresenta maior atrofia quando comparado ao hemisfério direito (LUBBEN et al., 2021). Tendo em vista essas informações sobre assimetrias e suas correlações com

a DA, a Variante Logopênica da Afasia Progressiva Primária (VLAPP) é considerada pelo o Departamento Científico de Neurologia Cognitiva e Envelhecimento¹ da Academia Brasileira de Neurologia² como uma característica para auxiliar o diagnóstico da DA. A VLAPP é caracterizada por alterações assimétricas que afetam majoritariamente regiões do hemisfério esquerdo (MONTEMBEAULT et al., 2018).

A inspeção visual das lesões utilizando biomarcadores é o método mais comum para o diagnóstico clínico da DA (SCHILLING et al., 2022). Entretanto, a inspeção visual de imagens é, inerentemente, subjetiva e pode ser enviesada pelos avaliadores (RAU; URBACH, 2021; LORETO et al., 2023). Entre os diversos métodos de inspeção visual disponíveis, o MTA é o mais utilizado (SCHILLING et al., 2022), porém, não há consenso quanto ao valor ideal do limiar que diferencia entre a neurodegeneração característica do envelhecimento e aquela decorrente da DA (CLAUS et al., 2017; RAU; URBACH, 2021; LORETO et al., 2023). Ademais, o MTA não é capaz de excluir outras causas de demências que também estão relacionadas com atrofia dessa região (RAU; URBACH, 2021; SCHILLING et al., 2022; LORETO et al., 2023).

Metodologias de análise de imagens são capazes de identificar corretamente a DA em imagens de RM, especialmente na distinção entre indivíduos CN e aqueles em estágios avançados da doença. Entretanto, a predição de conversão de indivíduos com Comprometimento Cognitivo Leve (CCL) para a DA permanece uma área em constante desenvolvimento e de grande interesse clínico. Nesse sentido, trabalhos recentes têm como objetivo classificar indivíduos do grupo CCL em CCL estável (CCLe), aqueles que permanecem com o diagnóstico, e CCL progressivo (CCLp), aqueles que progridem para DA (LONG; JIANG; ZHANG, 2018). A identificação precoce de indivíduos no grupo CCLp é fundamental para a aplicação de tratamentos que visam retardar a progressão dos sintomas. Dessa forma, o desenvolvimento de biomarcadores sensíveis às sutis alterações cerebrais que precedem o diagnóstico clínico definitivo é um dos principais desafios da pesquisa na área, motivando a exploração de novas fontes de informação contidas nas neuroimagens.

Somado a isso, a assimetria inter-hemisférica é um biomarcador promissor, porém ainda pouco explorado na predição específica da conversão para DA. A abordagem proposta neste estudo busca preencher essa lacuna ao quantificar as assimetrias inter-hemisféricas do encéfalo, com foco na região hipocampal. Para isso, a metodologia combina múltiplos tipos de informação: atributos volumétricos, que medem o tamanho das estruturas; descritores de forma geométrica, que capturam a sua morfologia tridimensional; e características de textura, que descrevem os padrões de intensidade da imagem. Esses atributos multimodais são então processados por uma arquitetura de aprendizado profundo, especificamente uma Rede Neural Siamesa, projetada para modelar e aprender a partir das

¹ <https://abneuro.org.br/dc/neurologia-cognitiva-e-envelhecimento>

² <https://abneuro.org.br>

diferenças sutis entre os dois hemisférios cerebrais.

Nesse contexto, o objetivo desta pesquisa foi o desenvolvimento de um novo método para quantificar assimetrias inter-hemisféricas nas populações CN, CCL e DA, visando tanto a classificação dos estágios da doença quanto, de forma crucial, a predição da conversão de indivíduos CCL para DA. Para isso, utilizou-se a combinação de técnicas de extração de atributos com aprendizado de máquina profundo em uma abordagem multimodal, buscando melhorar o entendimento sobre a influência das alterações inter-hemisféricas na progressão da doença. A contribuição deste trabalho reside na criação de um descritor unificado que captura de forma sensível as disparidades entre os hemisférios, com o potencial de se tornar uma ferramenta mais robusta para o prognóstico precoce, em linha com estudos que buscam novos biomarcadores baseados em imagem (LONG; JIANG; ZHANG, 2018; POLONI et al., 2021).

1.2 Objetivos

Os objetivos geral e específicos desta pesquisa são descritos a seguir.

1.2.1 Objetivo geral

O objetivo geral desta pesquisa foi o desenvolvimento de uma metodologia computacional com técnicas de visão computacional para classificar imagens de RM nas populações CN, CCL e DA, bem como à predição da conversão de indivíduos CCL para DA, por meio da diferenciação entre CCLe e CCLp, utilizando atributos inter-hemisféricos da região hipocampal.

1.2.2 Objetivos específicos

Para atingir o objetivo principal deste trabalho, foram estabelecidos os seguintes objetivos específicos:

- ❑ Organizar e pré-processar as imagens de RM das bases de dados IXI e ADNI.
- ❑ Detectar pontos salientes na região hipocampal de cada hemisfério cerebral utilizando a técnica Dual-Tree Complex Wavelet Transform (DT-CWT) e extrair seus respectivos descritores com a técnica *shape context*.
- ❑ Extrair atributos volumétricos da região hipocampal.
- ❑ Utilizar os atributos inter-hemisféricos com modelos de aprendizado de máquina para classificação binária entre os grupos: CN vs. DA, CN vs. CCL, CCL vs. DA e CCLe vs. CCLp.

- ❑ Comparar a significância estatística dos resultados obtidos entre os diferentes modelos propostos.
- ❑ Comparar os resultados obtidos com o estado da arte.

1.3 Organização do trabalho

O presente texto foi organizado em sete capítulos, incluindo esse capítulo introdutório. Os capítulos seguintes são compostos por:

- ❑ Capítulo 2: apresenta as informações técnicas sobre a DA, RM e os métodos utilizados na pesquisa.
- ❑ Capítulo 3: apresenta a revisão do estado da arte relacionada ao tópico de pesquisa principal desse trabalho.
- ❑ Capítulo 4: apresenta a proposta de pesquisa elaborada para atingir os objetivos desse trabalho.
- ❑ Capítulo 5: apresenta os experimentos e resultados obtidos.
- ❑ Capítulo 6: apresenta as conclusões e considerações para trabalhos futuros.

Capítulo 2

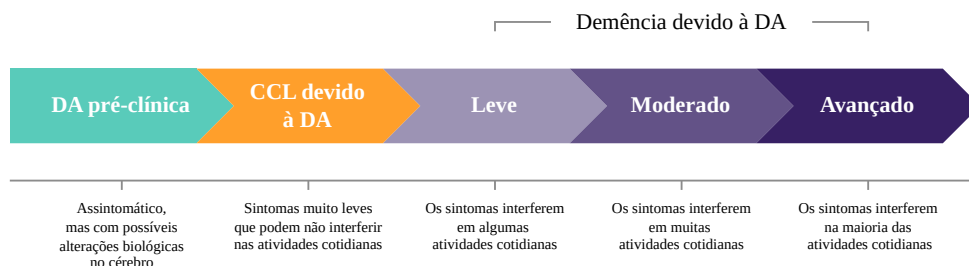
Fundamentação teórica

2.1 Doença de Alzheimer

A doença de Alzheimer é uma condição neurodegenerativa progressiva e irreversível, representando a causa mais comum de demência em todo o mundo, sendo responsável por 60 a 80% dos casos (Alzheimer's Association, 2023; JACK et al., 2024). A sua etiologia é multifatorial, envolvendo uma interação complexa entre fatores genéticos, ambientais e de estilo de vida (RASMUSSEN; LANGERMAN, 2019). Atualmente, a DA é definida como um processo biológico que se inicia com o surgimento de alterações neuropatológicas, mesmo em indivíduos assintomáticos, e cuja progressão leva ao aparecimento e agravamento dos sintomas clínicos (JACK et al., 2024).

O avanço da doença segue um contínuo que pode ser dividido em estágios, como ilustra a Figura 2.1. A fase pré-clínica, que pode durar até 20 anos, é caracterizada pela presença de biomarcadores da doença sem manifestações clínicas evidentes (Alzheimer's Association, 2023). A fase seguinte, o Comprometimento Cognitivo Leve, é um estágio intermediário em que o indivíduo apresenta alterações cognitivas que excedem o esperado para a idade, mas que não interferem significativamente nas atividades diárias. O CCL pode ser classificado como estável (CCLe), quando os sintomas permanecem constantes, ou progressivo, quando evolui para a demência. Por fim, a fase de demência é marcada por um declínio cognitivo e funcional severo, que impacta profundamente a autonomia do paciente (JACK et al., 2024).

Figura 1 – Contínuo da DA.



Fonte: Extraído, editado e traduzido de Alzheimer's Association (2023).

2.1.1 Causas e fisiopatologia

Do ponto de vista fisiopatológico, a DA é caracterizada por duas principais alterações neuropatológicas: o acúmulo de placas de peptídeo beta-amiloide ($A\beta$) no espaço extracelular e a formação de emaranhados neurofibrilares de proteína tau hiper-fosforilada no interior dos neurônios (JACK et al., 2024).

As placas amiloides resultam da clivagem anormal da proteína precursora de amiloide (do inglês, *Amyloid Precursor Protein* - APP). Esses agregados proteicos são tóxicos para os neurônios, interferindo na comunicação sináptica e desencadeando uma resposta inflamatória crônica que contribui para a morte celular (Alzheimer's Association, 2023).

A proteína tau, em seu estado normal, é responsável por estabilizar os microtúbulos, componentes essenciais do citoesqueleto neuronal. Na DA, a tau sofre hiperfosforilação, o que a leva a se desprender dos microtúbulos e a se agregar, formando os emaranhados neurofibrilares. Esse processo desestabiliza os microtúbulos, comprometendo o transporte de nutrientes e outras moléculas essenciais ao longo do axônio e, em última análise, levando à disfunção e morte do neurônio (JACK et al., 2024; DUBOIS et al., 2023).

Essas alterações patológicas resultam em uma atrofia cortical difusa, afetando inicialmente regiões como o hipocampo e o córtex entorrinal, cruciais para a memória, e expandindo-se progressivamente para outras áreas do cérebro, o que se correlaciona com a progressão dos sintomas (JACK et al., 2011; Alzheimer's Association, 2023; GAUTHIER et al., 2021a).

2.1.2 Sintomas e diagnóstico

Os sintomas da DA variam conforme o estágio da doença. Na fase inicial, são comuns a perda de memória para eventos recentes, dificuldade em encontrar palavras e desorientação espacial e temporal. Com a progressão, os déficits se acentuam, surgindo dificuldades no planejamento e execução de tarefas complexas, alterações de comportamento (apatia, agitação, depressão) e problemas de reconhecimento de familiares e amigos. Nos estágios avançados, o paciente torna-se totalmente dependente para as atividades de vida diária,

com comprometimento severo da linguagem e da motricidade (Alzheimer's Association, 2023).

O diagnóstico da DA é complexo e baseia-se na combinação de múltiplos critérios. A abordagem moderna, proposta por (JACK et al., 2024), enfatiza uma definição biológica da doença, onde a detecção de biomarcadores específicos é suficiente para o diagnóstico, independentemente da presença de sintomas. Os principais biomarcadores são divididos em:

- ❑ **Biomarcadores do peptídeo $A\beta$:** Medidos por Tomografia por Emissão de Pósitrons (do inglês, *Positron Emission Tomography* - PET) com ligantes específicos para amiloide ou pela análise do líquido cefalorraquidiano (*Cerebral Spinal Fluid* - CSF), onde se observa uma diminuição da concentração de $A\beta_{42}$.
- ❑ **Biomarcadores da patologia Tau:** Incluem a detecção de agregados de tau por PET (tau-PET) e a medição de formas de tau fosforilada no CSF ou, mais recentemente, no plasma sanguíneo.

A neuroimagem estrutural, especialmente a RM, desempenha um papel crucial ao revelar padrões de atrofia cerebral, como a redução do volume do hipocampo e do lobo temporal medial, que são característicos da DA (KNOPMAN et al., 2021; GAUTHIER et al., 2021b). Testes neuropsicológicos continuam sendo fundamentais para avaliar a extensão do comprometimento cognitivo e funcional do paciente (KNOPMAN et al., 2021; TAKIZAWA et al., 2015).

2.1.3 Tratamentos

Atualmente, não há cura para a DA. No entanto, os tratamentos disponíveis visam retardar a progressão dos sintomas e melhorar a qualidade de vida. As abordagens terapêuticas podem ser divididas em duas categorias principais (JACK et al., 2024):

- ❑ **Terapias Sintomáticas:** Incluem medicamentos como os inibidores da acetilcolinesterase (ex: donepezila, rivastigmina) e os antagonistas do receptor NMDA (memantina), que atuam regulando neurotransmissores e podem melhorar temporariamente a função cognitiva e os sintomas comportamentais.
- ❑ **Terapias Modificadoras da Doença:** Representam a fronteira da pesquisa em DA e têm como alvo a fisiopatologia subjacente da doença. Recentemente, foram aprovados os primeiros tratamentos com anticorpos monoclonais que visam remover as placas de beta-amiloide do cérebro (por exemplo, lecanemab e donanemab). Estudos demonstraram que essas terapias podem reduzir a carga amiloide e retardar modestamente o declínio cognitivo em pacientes nos estágios iniciais da doença.

2.2 Assimetrias inter-hemisféricas

O cérebro humano é caracterizado por uma notável assimetria, tanto estrutural quanto funcional, entre os hemisférios esquerdo e direito. Essa lateralização é um princípio fundamental da organização cerebral, acreditando-se que otimiza a eficiência do processamento neural e evita interferências entre funções concorrentes (OCKLENBURG et al., 2024). Funções como a linguagem e a praxia manual são predominantemente lateralizadas no hemisfério esquerdo na maioria dos indivíduos, enquanto a atenção visuoespacial e o processamento de emoções são predominantemente atribuídas ao hemisfério direito (OCKLENBURG et al., 2024).

Embora exista um padrão conhecido de organização hemisférica, há uma variabilidade interindividual substancial. Por exemplo, enquanto cerca de 95% dos destros apresentam dominância do hemisfério esquerdo para a linguagem, essa porcentagem cai para aproximadamente 70% nos canhotos (OCKLENBURG et al., 2024). Essa variabilidade é um fator crucial a ser considerado no contexto clínico.

2.2.1 Assimetria na apresentação e progressão da DA

Pesquisas recentes têm explorado a relação entre as assimetrias cerebrais e o desenvolvimento de doenças neurodegenerativas (OCKLENBURG et al., 2024; MATSUDA et al., 2025). A hipótese principal é que padrões atípicos de assimetria ou uma alteração na assimetria ao longo do tempo possam estar associados ao início ou à progressão dessas doenças.

A lateralidade dos sintomas é uma característica observada na DA. Pacientes podem apresentar predominantemente deficit de memória verbal e linguagem (sugerindo maior impacto no hemisfério esquerdo) ou deficit visuoespaciais (sugerindo maior impacto no hemisfério direito), dependendo de qual hemisfério apresenta maior carga de neurodegeneração e patologia amiloide e tau (OCKLENBURG et al., 2024; MATSUDA et al., 2025).

A investigação das assimetrias cerebrais na DA, portanto, oferece uma nova perspectiva para entender a heterogeneidade da doença. A análise de como as assimetrias estruturais e funcionais se alteram com a progressão da doença pode fornecer biomarcadores sensíveis para o diagnóstico precoce e monitoramento da progressão. A integração de medidas de assimetria com outros biomarcadores estabelecidos da DA é uma área promissora para a pesquisa futura.

2.3 Imagens de ressonância magnética

A imagem por RM é uma técnica de diagnóstico por imagem não invasiva que gera imagens transversais de alta resolução de estruturas internas sem o uso de radiação ionizante.

A RM explora o fenômeno da ressonância magnética nuclear, no qual os núcleos atômicos, quando expostos a um forte campo magnético, absorvem e reemitem ondas eletromagnéticas em uma frequência característica, ou ressonante, na faixa de Radiofrequência (RF) (YOUSAF; DERVENOULAS; POLITIS, 2018; MAZZOLA, 2009).

2.3.1 Princípios de formação da imagem

A formação da imagem em RM baseia-se na interação entre um forte campo magnético externo, pulsos de radiofrequência e os núcleos de hidrogênio (prótons) presentes abundantemente nos tecidos do corpo humano. Os prótons possuem uma propriedade quântica chamada spin, que, associada à sua carga elétrica positiva, gera um pequeno momento magnético, fazendo com que cada próton se comporte como um minúsculo ímã (MAZZOLA, 2009).

Em condições normais, esses momentos magnéticos estão orientados aleatoriamente, resultando em uma magnetização líquida nula. No entanto, quando um paciente é colocado dentro do magneto do equipamento de RM, os prótons se alinham com o forte campo magnético externo (B_0), seja de forma paralela (estado de menor energia) ou antiparalela (estado de maior energia). Como há um ligeiro excesso de prótons no estado de menor energia, surge uma magnetização líquida longitudinal, alinhada com a direção de B_0 (MAZZOLA, 2009; YOUSAF; DERVENOULAS; POLITIS, 2018).

Para gerar um sinal detectável, um pulso de radiofrequência (B_1) é aplicado perpendicularmente a B_0 . Se a frequência desse pulso corresponder à frequência de precessão dos prótons (frequência de Larmor), ocorre a ressonância. Os prótons absorvem energia e sua magnetização líquida é desviada para o plano transversal, pulso de excitação. Quando o pulso de RF é desligado, processo de relaxação, o sinal decai, liberando a energia absorvida e retornando ao seu alinhamento original com B_0 . Esse processo de relaxação gera um sinal de RF que é captado por uma bobina receptora e usado para construir a imagem (MAZZOLA, 2009).

Para localizar a origem do sinal no espaço e formar uma imagem, gradientes de campo magnético são aplicados em três direções ortogonais. Esses gradientes alteram ligeiramente o campo magnético em locais específicos, fazendo com que a frequência de precessão dos prótons varie de acordo com sua posição. Esse processo, conhecido como codificação espacial, permite mapear o sinal para coordenadas espaciais específicas, reconstruindo uma imagem detalhada da anatomia (YOUSAF; DERVENOULAS; POLITIS, 2018).

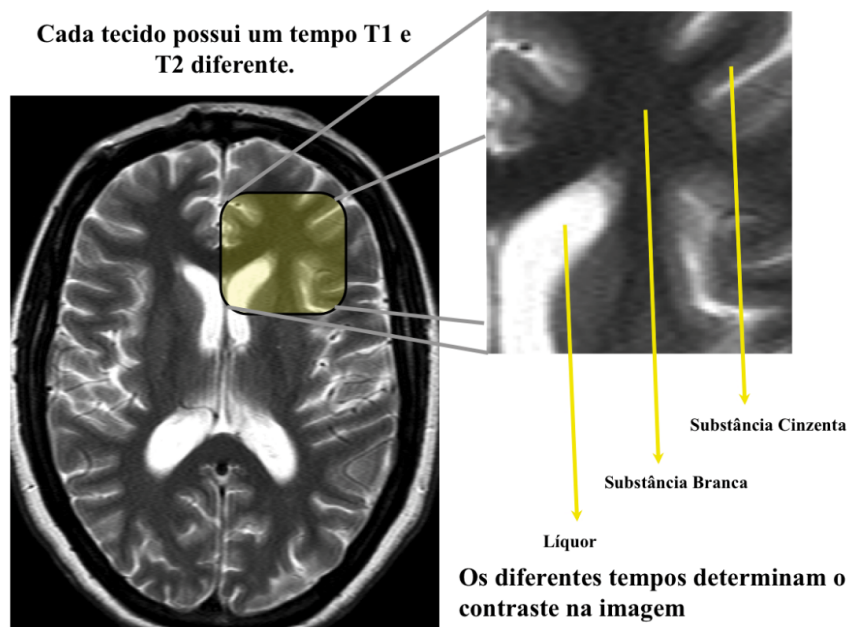
2.3.2 Contraste da imagem

O contraste entre diferentes tecidos na RM é determinado principalmente pelas taxas de relaxação dos prótons após o pulso de RF. Existem dois processos de relaxação principais:

- ❑ **Relaxação T1 (Longitudinal):** É o tempo que a magnetização líquida leva para retornar ao seu alinhamento com o campo magnético principal (B_0). Tecidos com T1 curto (ex: gordura) recuperam a magnetização rapidamente e aparecem brilhantes em imagens ponderadas em T1. Tecidos com T1 longo (ex: LCR) aparecem escuros.
- ❑ **Relaxação T2 (Transversal):** É o tempo que o sinal no plano transversal leva para decair devido à perda de coerência de fase entre os prótons. Tecidos com T2 longo (ex: CSF, patologias com edema) perdem a coerência lentamente e aparecem brilhantes em imagens ponderadas em T2, enquanto tecidos com T2 curto (ex: substância branca) aparecem escuros.

Ao manipular os parâmetros da sequência de pulsos, como o Tempo de Repetição (TR) e o Tempo de Eco (TE), é possível gerar imagens com diferentes ponderações (T1, T2, densidade de prótons), realçando características específicas dos tecidos e auxiliando no diagnóstico de diversas condições (MAZZOLA, 2009; YOUSAF; DERVENOULAS; POLITIS, 2018). A Figura 2.3.2 ilustra as diferenças nos tempos T2 entre os tecidos envolvidos.

Figura 2 – Imagem de RM ponderada em T2, mostrando na imagem ampliada os diferentes tecidos envolvidos nos tempos T2.



Fonte: Mazzola (2009).

2.4 Detecção do Plano Mediano

O alinhamento correto do encéfalo é um pré-processamento essencial para diversas análises de neuroimagem, como a avaliação de simetria e o registro de imagens. Neste

contexto, (FERRARI; PINTO; MOREIRA, 2016) desenvolveram uma técnica automática para a detecção do Plano Mediano ou Plano Médio-Sagital (*Mid-Sagittal Plane* - MSP) em imagens de RM. Diferentemente da maioria das abordagens que utilizam medidas de simetria entre os hemisférios cerebrais, esta técnica emprega uma medida direta para identificar as estruturas da linha média. O método utiliza uma medida de achatamento (*sheetness measure*), obtida a partir da análise de autovalores das respostas locais de congruência de fase 3D (*Phase Congruency* - PC), para determinar os *voxels* com maior probabilidade de pertencer à fissura longitudinal medial, que define o MSP.

Para determinar o plano de melhor ajuste que representa o MSP, um algoritmo de Weighted Least-Squares (WLS) é aplicado de forma iterativa, em uma abordagem progressiva de refinamento (*coarse-to-fine*). Inicialmente, um plano aproximado é estimado utilizando todos os *voxels* com valor de achatamento acima de um limiar, sem ponderação. Nas iterações seguintes, o ajuste é refinado em uma região de busca progressivamente menor, utilizando os valores da medida de achatamento como pesos para o algoritmo WLS, o que confere maior importância aos *voxels* com maior probabilidade de pertencer ao MSP. A técnica foi validada em 202 imagens de RM (162 sintéticas e 40 clínicas), demonstrando ser eficaz e robusta tanto para cérebros simétricos quanto assimétricos, com um erro angular médio absoluto inferior a $0,5^\circ$ em relação aos planos de referência.

2.5 Pré-processamento

O pré-processamento das imagens de RM é fundamental para restaurar e otimizar a qualidade das imagens, que são frequentemente afetadas por ruídos e distorções intrínsecas ao processo de aquisição. A formação de uma imagem de RM pode ser representada pela equação

$$v(x) = u(x)f(x) + n(x), \quad (1)$$

em que $v(x)$ simboliza o sinal amostrado em uma localização espacial x , u é o sinal emitido pelo tecido, f representa o efeito de degradação multiplicativo chamado de *bias field*, e n é o componente de ruído aditivo, presumido ser independente de u (SLED; ZIJDENBOS; EVANS, 1998). A partir desse modelo, é estabelecida uma metodologia de pré-processamento, inicialmente focada na eliminação do ruído aditivo, seguido pelo tratamento do *bias field* multiplicativo. Após a remoção desses elementos, é realizada a padronização de histogramas usando a técnica proposta por (NYÚL; UDUPA; ZHANG, 2000) e o alinhamento espacial das imagens de estudo com uma imagem de referência. Finalmente, define-se uma máscara para a remoção da caixa craniana.

Uma breve descrição de cada uma das etapas de pré-processamento citadas acima são apresentadas nas subseções seguintes.

2.5.1 Redução de ruído

O método NonLocal Means (NLM), originalmente proposto para lidar com ruído aditivo Gaussiano (BUADES; COLL; MOREL, 2005), foi adaptado com êxito para o contexto de imageamento por RM (COUPÉ et al., 2008). Este método baseia-se em uma média ponderada dos valores dos *voxels* dentro de uma janela de busca extensa, onde os *voxels* são ponderados com base em uma medida de similaridade. Essa similaridade é estimada pela diferença quadrada entre os *voxels* correspondentes dentro de janelas locais centralizadas em posições comparadas. O resultado é uma saída filtrada que preserva as estruturas da imagem enquanto suaviza o ruído. Ao analisar não apenas a vizinhança local de um *voxel*, mas várias regiões similares àquela vizinhança, consegue atenuar eficazmente o ruído Riciano em imagens de RM. O NLM é particularmente eficiente em imagens com alta relação sinal-ruído Signal-to-Noise Ratio (SNR), onde a distribuição do ruído se aproxima da Gaussiana (COUPÉ et al., 2008).

2.5.2 Correção do efeito *bias field*

O efeito *bias field* Bias Field (BF) é caracterizado por variações nos níveis de intensidade da imagem que podem ser causadas por flutuações no campo magnético dentro do *scanner*, variações nas características da bobina de recepção e pela interação entre o campo magnético e o paciente (BELAROUSSI et al., 2006; VOVK; PERCUS; LIKAR, 2007). O algoritmo N4-ITK (TUSTISON et al., 2010), uma evolução do Nonparametric Nonuniform Intensity Normalization (N3) proposto por Sled, Zijdenbos e Evans (1998), é utilizado para corrigir o campo de heterogeneidades nas imagens 3D de RM. O N4-ITK utiliza um aproximador *B-spline* mais estável e robusto a ruídos, e um esquema de otimização hierárquico modificado. Este algoritmo adota uma abordagem de regularização baseada em *B-splines* e realiza otimizações sucessivas para refinar de forma iterativa a estimação do *bias field*. Além disso, o N4-ITK emprega uma estratégia de suavização mais eficaz que melhora a performance de convergência. A implementação usada nesse trabalho está disponível na biblioteca ITK¹.

2.5.3 Normalização de intensidade

A normalização de intensidades foi realizada utilizando o método baseado em Fuzzy C-Means (FCM) (REINHOLD et al., 2019). Este método emprega o algoritmo FCM para calcular uma máscara de substância branca (SB) na imagem. A máscara de SB é então utilizada para normalizar a imagem inteira para a média da intensidade da substância branca. O procedimento envolve a determinação da média da substância branca (u_{fcm}) dentro da máscara W , que é um subconjunto da máscara cerebral B da imagem I , utilizando a fórmula:

¹ Disponível em: <<https://itk.org/>> Acesso em 01 dez, 2023

$$\mu_{fcm} = \frac{1}{|W|} \sum_{w \in W} I(w).$$

A imagem normalizada por FCM, $I_{fcm}(x)$, é então calculada como:

$$I_{fcm}(x) = \frac{c_1 \cdot I(x)}{\mu_{fcm}},$$

onde c_1 é uma constante positiva que define a média da substância branca após a normalização.

2.5.4 Alinhamento espacial

Em imagens médicas, o corregristo de imagens é um processo de alinhamento espacial de duas ou mais imagens de um mesmo órgão anatômico, obtidas sob diferentes condições ou momentos. O intuito dessa etapa é alinhar todas as imagens em estudo num único sistema de coordenadas, garantindo assim a consistência espacial nas análises subsequentes. Nesse trabalho, os processos de corregristo foram realizados com a ferramenta de código aberto Elastix² desenvolvida nos estudos de Klein et al. (2010) e Shamonin et al. (2014).

No processo de corregristo, uma imagem móvel, $I_M(\mathbf{x})$, é deformada e ajustada sobre a imagem fixa, $I_F(\mathbf{x})$, por meio de uma transformação de coordenadas $T_\mu(\mathbf{x})$, tal que $I_M(T(\mathbf{x}))$ fique espacialmente alinhada à $I_F(\mathbf{x})$. As imagens fixa e móvel possuem dimensão d e cada uma possui seu próprio domínio espacial: $\Omega_F \subset \mathbb{R}^d$ e $\Omega_M \subset \mathbb{R}^d$, respectivamente (KLEIN et al., 2009).

O modelo de transformação $T_\mu(\mathbf{x})$ usado determina os tipos de deformações aplicadas à imagem móvel para alinhá-la à imagem fixa. Essa transformação é definida como:

$$T_\mu(\mathbf{x}) = A(\mathbf{x} - \mathbf{t}) + \mathbf{t} + \mathbf{c}, \quad (2)$$

sendo $\mathbf{t}$ o vetor de translação, $\mathbf{c}$ o centro de rotação, e A uma matriz que contém o fator de escalonamento, rotação e cisalhamento. O vetor paramétrico μ é formado com os elementos a_{ij} da matriz A e vetor de translação $\mathbf{t}$. Em duas dimensões, μ possui seis elementos: $(a_{11}, a_{12}, a_{21}, a_{22}, t_x, t_y)^T$, e doze elementos em três dimensões (KLEIN et al., 2010).

O corregristo deformável *B-spline*, definido como a parametrização de

$$T_\mu(\mathbf{x}) = \mathbf{x} + \sum_{\mathbf{x}_k \in \mathcal{N}_x} p_k \beta^3 \left(\frac{\mathbf{x} - \mathbf{x}_k}{\sigma} \right), \quad (3)$$

com $\mathbf{x}_k$ sendo os pontos de controle, $\beta^3(\mathbf{x})$ o polinômio *B-spline*, p_k representando os vetores de coeficientes, σ o espaçamento entre os pontos de controle, e $\mathcal{N}_x$ sendo o conjunto de todos os pontos de controle que são afetados pela transformação *B-spline* em $\mathbf{x}$ (KLEIN et al., 2010).

² Disponível em: <<https://elastix.lumc.nl/>> Acesso em: 01 dez, 2023.

A principal tarefa do corregristo é encontrar a transformação de coordenadas T_{μ} , sendo esse processo formulado matematicamente como um problema de otimização, tal que:

$$\hat{\mu} = \arg \min_{\mu} \mathcal{C}(T_{\mu}; I_F, I_M), \quad (4)$$

em que μ corresponde ao vetor de parâmetros que modela a transformação T_{μ} , e $\mathcal{C}$ é uma função de custo que mede a similaridade do alinhamento entre as imagens. Neste caso, $\hat{\mu}$ corresponde ao vetor de parâmetros μ atualizado pelo processo de otimização.

Além da transformação e função de custo, o software Elastix disponibiliza outros componentes que podem ser escolhidos, como os componentes de otimização, estratégias de amostragem, interpoladores e estratégias hierárquicas (pirâmides de imagens). Os parâmetros usados por esses componentes são escolhidos *a priori* por meio de um arquivo de texto usado como entrada pelo algoritmo. Neste trabalho, foram utilizados os parâmetros *par0010affine.txt* e *par0010bspline.txt* para as transformações afim e *B-spline*, disponíveis no *Elastix Model Zoo*³.

2.5.5 Extração do encéfalo

A extração do encéfalo, ou remoção do crânio (do inglês, skull-stripping), é uma etapa de pré-processamento que visa segmentar e isolar o tecido cerebral de estruturas adjacentes, como o crânio e glóbulos oculares. A abordagem SynthStrip (HOOPES et al., 2022) introduz uma metodologia que emprega uma rede neural convolucional treinada exclusivamente com um conjunto de dados sintéticos, tornando-a agnóstica ao contraste da imagem de entrada. Os métodos tradicionais de extração do encéfalo são comumente otimizados com atributos específicos, como alta resolução isotrópica e contraste T1-*w*, o que limita a aplicabilidade em diferentes dados clínicos que exibem variabilidade em contraste, resolução e artefatos.

Para mitigar as deficiências das abordagens convencionais, o SynthStrip emprega uma estratégia treinamento com dados sintéticos, gerados a partir de segmentações anatômicas pré-existentes. Este processo de síntese permitiu a criação de imagens com uma gama deliberadamente ampla de anatomias, distribuições de intensidade e artefatos, superando a diversidade encontrada em dados médicos reais. Ao expor uma rede neural convolucional a este vasto espectro de variações, o SynthStrip adquire a capacidade de generalizar para diferentes modalidades de RM. Os resultados apresentados demonstram que o SynthStrip alcança uma precisão superior em comparação com métodos estabelecidos em um conjunto diversificado de imagens.

³ Disponível em: <<https://elastix.lumc.nl/modelzoo/par0010/>> Acesso em: 01 maio, 2023.

2.6 Transformadas Wavelets

A transformada *wavelet* é uma ferramenta utilizada em diversas técnicas de processamento de imagens e sinais (SELESNICK; BARANIUK; KINGSBURY, 2005). Esta seção descreve a DT-CWT e as transformadas *wavelet* que a antecederam, abordando-as motivações de cada técnica e suas distinções.

2.6.1 Transformada *wavelet* discreta (DWT)

A transformada Transformada Wavelet Discreta (DWT) é uma técnica de processamento de sinais que substitui a transformada de Fourier, que contém infinitas funções periódicas, por um conjunto de funções localizadas no tempo e com suporte compacto, denominadas *wavelets* (SELESNICK; BARANIUK; KINGSBURY, 2005). O termo "discreto" relativo à transformada, refere-se à discretização dos parâmetros das *wavelets* filhas e não à discretização do tempo. Na forma contínua, um sinal $x(t)$ pode ser decomposto em termos de *wavelets* e funções de escalonamento como:

$$x(t) = \sum_{n=-\infty}^{\infty} c(n)\phi(t-n) + \sum_{j=0}^{\infty} \sum_{n=-\infty}^{\infty} d(j,n)2^{j/2}\psi(2^j t - n), \quad (5)$$

em que c e d representam respectivamente, os coeficientes de escala e coeficientes *wavelets*, que são computados pelos produtos internos:

$$c(n) = \int_{-\infty}^{\infty} f(t)\phi(t-n)dt \quad (6)$$

$$d(s,n) = 2^{s/2} \int_{-\infty}^{\infty} f(t)\psi(2^s t - n)dt. \quad (7)$$

Embora, na sua forma discreta, a DWT é amplamente usada para compressão de imagens (como o padrão JPEG 2000), ela apresenta quatro limitações: (i) falta de invariância à deslocamentos, (ii) baixa seletividade direcional, (iii) oscilações e (iv) *aliasing* (SELESNICK; BARANIUK; KINGSBURY, 2005). Na DWT, pequenos deslocamentos no tempo do sinal de entrada podem causar grandes variações nas distribuições de energia entre os coeficientes da DWT em diferentes escalas (KINGSBURY, 1998; BRASSAROTE; SOUZA; MONICO, 2018). A segunda limitação ocorre nos casos onde seria necessário diferenciar orientações na diagonal, nesse caso tanto a DWT quanto sua versão invariante ao deslocamento (*Discrete Wavelet Packets Transform* - DWPT), não são capazes de diferenciar as orientações em sentidos opostos, e.g., as orientações $\pm 45^\circ$ (KINGSBURY, 1998; NAFORNITA; ISAR; NAFORNITA, 2012). Os coeficientes da transformada *wavelet* funcionam como funções passa banda, oscilando entre valores positivos e negativos. Nesse sentido, os coeficientes podem apresentar valores baixos, ou até nulos, quando aplicados para filtrar singularidades. O efeito de *aliasing*, que ocorre quando a taxa de amostragem não é suficiente para reconstruir uma imagem, está presente na DWT pois os múltiplos

downsamplings necessários para calcular os coeficientes da transformada, podem apresentar as limitações citadas e não representar adequadamente os filtros necessários para reconstruir esse sinal (SELESNICK; BARANIUK; KINGSBURY, 2005).

2.6.2 Transformada *wavelet* complexa (CWT)

A transformada *wavelet* complexa (do inglês, Transformada Wavelet Complexa (CWT)) utiliza uma abordagem similar à transformada de Fourier para solucionar os problemas da DWT contínua. Isto é, diferentemente da DWT, que é baseada em *wavelets* oscilatórias com bases reais, a transformada de Fourier utiliza funções sinusoidais com componentes complexos, ou seja, $e^{i\Omega t} = \cos(\Omega t) + j\sin(\Omega t)$, em que $j = \sqrt{-1}$. O par definido pelos componentes reais e imaginários formam uma transformada Hilbert ($\mathcal{H}$). Assim, o sinal analítico $e^{i\Omega t}$ é representado apenas com metade da frequência do espectro (SELESNICK; BARANIUK; KINGSBURY, 2005). Dessa forma, a função *wavelet* complexa é definida como

$$\psi_c(t) = \psi_r(t) + j\psi_i(t), \quad (8)$$

em que $\psi_r(t)$ é uma função par no domínio real e $j\psi_i(t)$ é uma função ímpar com o componente imaginário. Esse sinal, projetado em $2^{j/2}\psi_c(2^j t - n)$ resulta no coeficiente *wavelet* complexo

$$d_c(j, n) = d_r(j, n) + jd_i(j, n), \quad (9)$$

com magnitude

$$|d_c(j, n)| = \sqrt{[d_r(j, n)]^2 + [d_i(j, n)]^2}, \quad (10)$$

e fase

$$\angle d_c(j, n) = \arctan\left(\frac{d_i(j, n)}{d_r(j, n)}\right), \quad (11)$$

quando $|d_c(j, n)| > 0$ (SELESNICK; BARANIUK; KINGSBURY, 2005). A CWT possibilita novos algoritmos de processamento de sinal multiescala que exploram a magnitude e a fase do sinal. Assim, uma grande magnitude indica a presença de uma singularidade, enquanto a fase indica sua posição dentro do suporte da *wavelet*.

2.6.3 Transformada *wavelet* complexa de árvore dupla bidimensional (DT-CWT 2D)

Com o objetivo de solucionar os problemas da DWT aplicada a imagens, Kingsbury (1998) propuseram a transformada *wavelet* complexa de árvore dupla (*Dual-Tree Complex Wavelet Transform* - DT-CWT) bidimensional, que utiliza uma representação em árvore com duas transformadas do tipo DWT. A primeira transformada representa a parte real enquanto a segunda representa a parte imaginária (SELESNICK; BARANIUK; KINGSBURY, 2005). Isto é, seja $\psi(x, y) = \psi(x)\psi(y)$ uma *wavelet* 2D separável,

ou seja, aplicada separadamente ao longo das linhas e colunas via *wavelets* 1D, em que $\psi(x) = \psi_h(x) + j\psi_g(x)$, então $\psi(x, y)$ pode ser calculada como

$$\begin{aligned}\psi(x, y) &= [\psi_h(x) + j\psi_g(x)] [\psi_h(y) + j\psi_g(y)] \\ &= \psi_h(x)\psi_h(y) - \psi_g(x)\psi_g(y) + j[\psi_g(x)\psi_h(y) + \psi_h(x)\psi_g(y)].\end{aligned}\quad (12)$$

A Equação 12 pode ser decomposta em partes reais e imaginárias, onde cada parte continua é uma *wavelet* complexa separável por duas *wavelets*. Nesse sentido, essas *wavelets* são definidas como

$$\text{Parte Real } \{\psi(x, y)\} = \psi_h(x)\psi_h(y) - \psi_g(x)\psi_g(y) \quad (13)$$

$$\text{Parte Imaginária } \{\psi(x, y)\} = \psi_g(x)\psi_h(y) + \psi_h(x)\psi_g(y). \quad (14)$$

Considerando as partes imaginárias $\psi(x)\overline{\psi(y)}$, $\phi(x)\psi(y)$, $\psi(x)\phi(y)$, $\phi(x)\overline{\psi(y)}$, e $\psi(x)\overline{\phi(y)}$, em que $\overline{f(\cdot)}$ representa o conjugado complexo de $f(\cdot)$, e

$$\psi(x) = \psi_h(x) + j\psi_g(x) \quad (15)$$

$$\phi(x) = \phi_h(x) + j\phi_g(x). \quad (16)$$

As seis sub-bandas da transformada que decompõem o sinal nas direções $\pm 15^\circ$, $\pm 45^\circ$ e $\pm 75^\circ$, são definidas como

$$\psi_i(x, y) = \frac{1}{\sqrt{2}} (\psi_{1,i}(x, y) + \psi_{2,i}(x, y)) \quad (17)$$

$$\psi_{i+3}(x, y) = \frac{1}{\sqrt{2}} (\psi_{1,i}(x, y) - \psi_{2,i}(x, y)), \quad (18)$$

para $i = 1, 2, 3$ (SELESNICK; BARANIUK; KINGSBURY, 2005). Uma ilustração das *wavelets* resultantes é apresentada na Figura 2.6.3. As *wavelets* separáveis 2D utilizadas como base nesse calculado são definidas pela combinação das DWT separáveis:

$$\psi_{1,1}(x, y) = \phi_h(x)\psi_h(y), \quad \psi_{2,1}(x, y) = \phi_g(x)\psi_g(y), \quad (19)$$

$$\psi_{1,2}(x, y) = \psi_h(x)\phi_h(y), \quad \psi_{2,2}(x, y) = \psi_g(x)\phi_g(y), \quad (20)$$

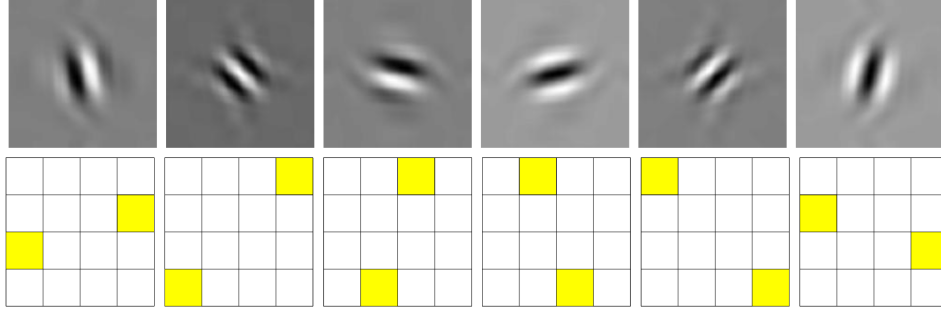
$$\psi_{1,3}(x, y) = \psi_h(x)\psi_h(y), \quad \psi_{2,3}(x, y) = \psi_g(x)\psi_g(y). \quad (21)$$

2.6.4 Transformada *wavelet* complexa de árvore dupla tridimensional (DT-CWT 3D)

De maneira análoga à implementação em duas dimensões, se a *wavelet* complexa $\psi(\cdot)$ é definida como $\psi(x) = \psi_h(x) + j\psi_g(x)$, então a *wavelet* complexa 3D é definida como

$$\begin{aligned}\psi(x, y, z) &= \psi(x)\psi(y)\psi(z) \\ &= (\psi_h(x) + j\psi_g(x)) (\psi_h(y) + j\psi_g(y)) (\psi_h(z) + j\psi_g(z)),\end{aligned}\quad (22)$$

Figura 3 – Representação das *wavelets* no domínio do espaço (primeira linha) e o suporte das *wavelets* do espectro de Fourier.



Fonte: Selesnick, Baraniuk e Kingsbury (2005).

em que $\psi_g(t)$ representa, aproximadamente, a transformada de Hilbert $\psi_h(t)$, $[\psi_g(t) \approx \mathcal{H}\{\psi_h(t)\}]$ (SELESNICK; LI, 2003). O componente real de $\psi(x, y, z)$ é utilizado para obter a *wavelet* 3D orientada:

$$\text{Parte Real}\{\psi(x, y, z)\} = \psi_1(x, y, z) - \psi_2(x, y, z) - \psi_3(x, y, z) - \psi_4(x, y, z), \quad (23)$$

sendo que

$$\psi_1(x, y, z) := \psi_h(x)\psi_h(y)\psi_h(z) \quad (24)$$

$$\psi_2(x, y, z) := \psi_g(x)\psi_g(y)\psi_h(z) \quad (25)$$

$$\psi_3(x, y, z) := \psi_g(x)\psi_h(y)\psi_g(z) \quad (26)$$

$$\psi_4(x, y, z) := \psi_h(x)\psi_g(y)\psi_g(z). \quad (27)$$

Em três dimensões, é necessário combinar quatro transformadas *wavelets* separáveis, utilizando as partes reais de $\psi(x)\psi(y)\overline{\psi(z)}$, $\psi(x)\overline{\psi(y)}\psi(z)$ e $\psi(x)\overline{\psi(y)}\overline{\psi(z)}$. A combinação da matriz ortonormal resultante são quatro transformadas *wavelets* 3D separáveis (SELESNICK; LI, 2003), definidas como

$$\psi_a(x, y, z) = 0,5 (\psi_1(x, y, z) - \psi_2(x, y, z) - \psi_3(x, y, z) - \psi_4(x, y, z)) \quad (28)$$

$$\psi_b(x, y, z) = 0,5 (\psi_1(x, y, z) - \psi_2(x, y, z) + \psi_3(x, y, z) + \psi_4(x, y, z)) \quad (29)$$

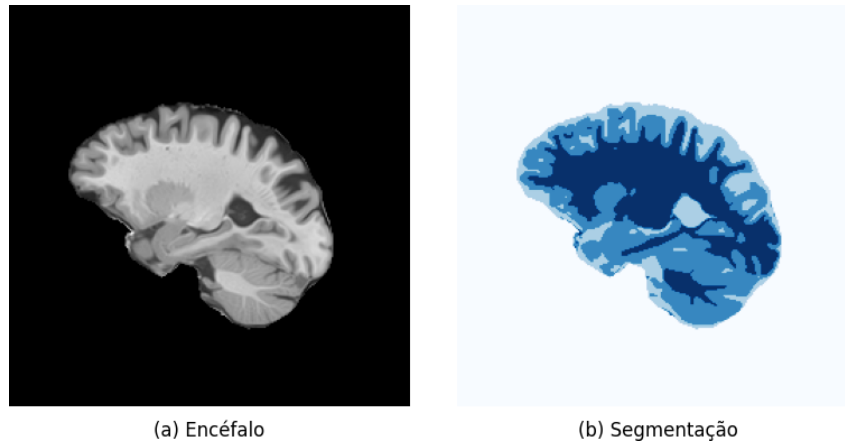
$$\psi_c(x, y, z) = 0,5 (\psi_1(x, y, z) + \psi_2(x, y, z) - \psi_3(x, y, z) + \psi_4(x, y, z)) \quad (30)$$

$$\psi_d(x, y, z) = 0,5 (\psi_1(x, y, z) + \psi_2(x, y, z) + \psi_3(x, y, z) - \psi_4(x, y, z)). \quad (31)$$

2.7 Segmentação de tecidos do encéfalo

A segmentação de imagens de ressonância magnética do encéfalo em diferentes classes de tecido, como substância cinzenta (*Grey Matter* - GM), substância branca (*White Matter* - WM) e líquido cefalorraquidiano (CSF), é uma etapa fundamental em neuroimagem para a análise quantitativa de estruturas cerebrais. Dentre as diversas técnicas desenvolvidas para esta finalidade, os métodos baseados em modelos estatísticos e probabilísticos

Figura 4 – Exemplo de segmentação de imagem de RM do encéfalo em CSF, GM, WM e *background*.



Fonte: elaborado pelo autor.

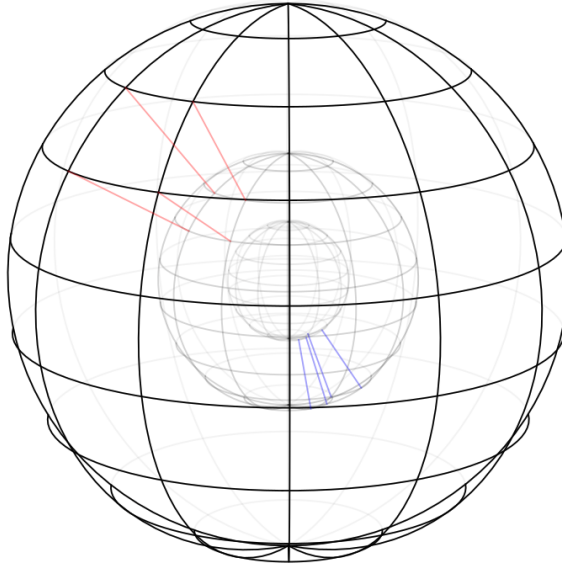
se destacam pela sua robustez e acurácia. Uma dessas técnicas é o algoritmo Atropos (AVANTS et al., 2011; WANG; YUSHKEVICH, 2013), uma ferramenta de código aberto que implementa um método de segmentação multivariada de n-classes.

O Atropos utiliza uma abordagem Bayesiana para o problema de segmentação, que é solucionado através do algoritmo Expectation-Maximization (EM). A formulação Bayesiana permite a incorporação de conhecimentos prévios sobre a distribuição espacial dos tecidos, utilizando mapas de probabilidade a priori, e modela as intensidades de cada classe de tecido por meio de misturas finitas paramétricas ou não paramétricas. Além disso, o algoritmo pode empregar modelos de Markov Random Field (MRF) para garantir a coerência espacial da segmentação, favorecendo configurações em que voxels vizinhos pertençam à mesma classe tecidual. Essa combinação de modelagem de intensidade, priors espaciais e regularização via MRF confere ao Atropos grande flexibilidade e precisão na delimitação das estruturas encefálicas.

A implementação do Atropos, distribuída com o pacote de software Advanced Normalization Tools (ANTs)⁴, é eficiente em termos de uso de memória, o que a torna capaz de lidar com um grande número de classes de segmentação. O método foi validado em diversos conjuntos de dados públicos, como o BrainWeb, demonstrando desempenho competitivo em tarefas de segmentação de três tecidos (GM, WM e CSF). A capacidade de integrar dados multimodais (e.g., imagens ponderadas em T1 e T2) e a interação com ferramentas de correção de inhomogeneidade de intensidade, como o N4, reforçam a sua aplicabilidade em uma vasta gama de estudos clínicos e de pesquisa que necessitem de uma segmentação encefálica precisa e reprodutível.

⁴ <<https://github.com/ANTsX/ANTs>>

Figura 5 – Histograma log-esférico utilizado para calcular o descritor *shape context*.



Fonte: elaborado pelo autor.

2.8 Shape context

Belongie, Malik e Puzicha (2002) introduziram o descritor Shape Context (SC) com o objetivo de extrair um vetor de características para o cálculo de correspondências entre objetos em imagens 2D. Nessa técnica, um objeto é representado por um conjunto finito $\mathcal{P} = \{p_1, \dots, p_n\}$, $p_i \in \mathbb{R}^d$ de n pontos discretos d dimensionais, que podem estar presentes no contorno interno ou externo do objeto, não sendo obrigatoriamente associados aos pontos de máxima curvatura ou inflexão. (BELONGIE; MALIK; PUZICHA, 2002). Para cada ponto $p_i \in \mathcal{P}$, calcula-se um histograma log-polar h_i das coordenadas relativas do ponto aos demais $n - 1$ pontos, sendo esse histograma chamado de *shape context* de p_i (BELONGIE; MALIK; PUZICHA, 2002). Esse descritor é, por definição, invariante à translação, uma vez que todas as medições utilizam pontos no contorno da forma, e para obter invariância à rotação, o histograma pode ser ajustado levando-se em consideração a orientação local (BELONGIE; MALIK; PUZICHA, 2002).

Frome et al. (2004) expandiram a técnica para dados tridimensionais utilizando um histograma log-esférico, ilustrado na Figura 5. Formalmente, para um determinado ponto $\mathbf{x} = (x, y, z)$ e uma vizinhança de raio R_{max} em $\mathbf{x}$, a técnica SC divide a vizinhança em n_r distâncias radiais, n_a ângulos azimute e n_e ângulos de elevação, e calcula os *bins* do

histograma com as seguintes fórmulas:

$$b_r = \max \left(0, n_r + \lfloor \log_2 \left(\frac{\|\mathbf{x} - \mathbf{x}_0\|}{R_{max}} \right) \rfloor \right), \quad (32)$$

$$b_a = \text{mod} \left(\left\lfloor \frac{n_a(\pi + \phi)}{2\pi} \right\rfloor, n_a \right), \quad (33)$$

$$b_e = \begin{cases} \text{mod} \left(\left\lfloor \frac{n_e(\frac{\pi}{2} + \theta)}{\pi} \right\rfloor, n_e \right) & \text{se } -\frac{\pi}{2} \leq \theta < \frac{\pi}{2} \\ n_e - 1 & \text{se } \theta = \frac{\pi}{2} \end{cases}, \quad (34)$$

$$b = b_r + n_r(b_a + n_a b_e), \quad (35)$$

em que b representa o índice no vetor 1D do histograma, e θ e ϕ representam, respectivamente, o ângulo de elevação e ângulo azimutal relativos, e são definidos como

$$\phi = \arctan \left(\frac{j - j_0}{i - i_0} \right), \quad (36)$$

$$\theta = \arctan \left(\frac{k - k_0}{\sqrt{(i - i_0)^2 + (j - j_0)^2}} \right), \quad (37)$$

com incrementos radiais no intervalo $R_{max}/2^{n_r-1}, R_{max}/2^{n_r-2}, \dots, R_{max}/2$, e *bins* igualmente divididos no intervalo $\phi \in [-\pi, \pi]$ e $\theta \in [-\pi/2, \pi/2]$ (SOUZA; POLONI; FERRARI, 2020).

2.8.1 Correspondências entre descritores

Seja p_i e q_j pontos em diferentes objetos, a função $C_{ij} = C(p_i, q_j)$ representa o custo de correspondência entre esses dois pontos. Na representação com descritores SC, a correspondência entre esses pontos é definida pelo teste χ^2 definido como

$$C_{ij} \equiv C(p_i, q_j) = \frac{1}{2} \sum_{k=1}^K \frac{[h_i(k) - h_j(k)]^2}{h_i(k) + h_j(k)}, \quad (38)$$

em que $h_i(k)$ e $h_j(k)$ representam o k -ésimo *bin* do histograma normalizado, para o ponto p_i e q_j , respectivamente (BELONGIE; MALIK; PUZICHA, 2002).

A distância χ^2 quantifica a dissimilaridade entre dois descritores, no entanto, para quantificar esse valor entre todos os descritores e obter os melhores casamentos, pode ser utilizada a minimização do custo C_{ij} entre todos os pares de pontos de $p_i \in \mathcal{P}$ e $q_j \in \mathcal{Q}$. Formalmente, deseja-se minimizar o custo total definido como

$$H(\pi) = \sum_i C(p_i, q_{\pi(i)}), \quad (39)$$

em que π representa a permutação com representação binária, *one-to-one*, entre os elementos de cada conjunto (BELONGIE; MALIK; PUZICHA, 2002). Essa minimização pode ser compreendida como um problema de emparelhamento máximo em grafos bipartidos, e resolvida em tempo $O(N^3)$ com o método Húngaro (JONKER; VOLGENANT,

1987), sendo a entrada do algoritmo a matriz quadrada C_{ij} (BELONGIE; MALIK; PUZICHA, 2002). Na prática, nos casos em que $|\mathcal{P}| \neq |\mathcal{Q}|$ a matriz resultante não será quadrada, podendo ser concatenada com valores *dummy* para permitir o cálculo com o método Húngaro (BELONGIE; MALIK; PUZICHA, 2002).

2.9 XGBoost

O XGBoost (eXtreme Gradient Boosting) é um algoritmo de aprendizado de máquina altamente eficaz e amplamente utilizado, baseado na técnica de *tree boosting* (potenciação de árvores). Desenvolvido por CHEN e GUESTRIN (2016), o XGBoost é conhecido por seu desempenho de ponta em uma variedade de tarefas de classificação, regressão e ranqueamento, dominando muitas competições de ciência de dados. Sua popularidade deriva não apenas de sua precisão, mas também de sua escalabilidade e eficiência computacional, sendo capaz de processar bilhões de exemplos com recursos significativamente menores do que outros sistemas (CHEN; GUESTRIN, 2016).

A essência do XGBoost reside em sua abordagem de construir um modelo de previsão na forma de um *ensemble* de árvores de decisão fracas, tipicamente árvores de regressão e classificação (CART). Diferente de algoritmos como o Random Forest, que constrói árvores de forma independente e paralela, o XGBoost constrói as árvores de maneira sequencial e aditiva. Cada nova árvore é treinada para corrigir os erros residuais cometidos pelo conjunto de árvores já existentes, melhorando iterativamente a precisão do modelo final.

2.9.1 Gradient Tree Boosting

O *Gradient Tree Boosting* é o princípio fundamental por trás do XGBoost. Em um modelo de *ensemble* de árvores, a predição final para uma dada amostra x_i é a soma das predições de K árvores de decisão individuais:

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), \quad f_k \in \mathcal{F} \quad (40)$$

em que $\mathcal{F}$ é o espaço de todas as possíveis árvores de decisão (CARTs), e cada f_k representa uma árvore específica com sua própria estrutura e pesos nas folhas. O treinamento desse modelo é feito de forma aditiva. Em cada iteração t , uma nova árvore f_t é adicionada ao *ensemble* para minimizar a função de objetivo, que geralmente consiste em uma função de perda l e um termo de regularização Ω :

$$\mathcal{L}^{(t)} = \sum_{i=1}^n l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) \quad (41)$$

O *Gradient Boosting* utiliza o gradiente da função de perda para guiar o treinamento da nova árvore. Em vez de ajustar a árvore f_t para prever diretamente o resíduo $y_i -$

$\hat{y}_i^{(t-1)}$, ela é ajustada para prever o gradiente negativo da função de perda em relação à predição atual. O XGBoost aprimora essa ideia utilizando não apenas a primeira derivada (gradiente), mas também a segunda derivada (Hessiana) da função de perda. Isso é feito através de uma aproximação de Taylor de segunda ordem da função de objetivo, o que permite um passo de otimização mais preciso e uma convergência mais rápida (CHEN; GUESTRIN, 2016).

2.9.2 O Algoritmo XGBoost e sua Função de Objetivo Regularizada

Uma das principais inovações do XGBoost é sua função de objetivo regularizada, que penaliza a complexidade do modelo para evitar o *overfitting*. A função de objetivo geral é:

$$\mathcal{L}(\phi) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \quad (42)$$

em que o termo de regularização Ω é definido para cada árvore f como:

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \|w\|^2 \quad (43)$$

Aqui, T é o número de folhas na árvore, w é o vetor de pesos (pontuações) das folhas, e γ e λ são hiperparâmetros que controlam a força da regularização. O termo γT penaliza árvores com muitas folhas, incentivando estruturas mais simples, enquanto o termo $\frac{1}{2} \lambda \|w\|^2$ (regularização L2) penaliza pesos de folha com magnitude grande, tornando as predições mais suaves e menos sensíveis a exemplos individuais (CHEN; GUESTRIN, 2016).

Ao otimizar a função de objetivo em cada passo aditivo usando a aproximação de Taylor de segunda ordem, o XGBoost deriva uma pontuação de qualidade para uma determinada estrutura de árvore q . Essa pontuação permite avaliar quão boa é uma divisão potencial em um nó da árvore. A redução de perda após uma divisão é calculada como:

$$\mathcal{L}_{split} = \frac{1}{2} \left[\frac{(\sum_{i \in I_L} g_i)^2}{\sum_{i \in I_L} h_i + \lambda} + \frac{(\sum_{i \in I_R} g_i)^2}{\sum_{i \in I_R} h_i + \lambda} - \frac{(\sum_{i \in I} g_i)^2}{\sum_{i \in I} h_i + \lambda} \right] - \gamma \quad (44)$$

onde g_i e h_i são a primeira e a segunda derivadas da função de perda, respectivamente, e I_L , I_R e I são os conjuntos de instâncias nos nós esquerdo, direito e pai. O algoritmo busca avidamente a divisão que maximiza essa pontuação, construindo a árvore de forma eficiente.

2.9.3 Otimizações e Escalabilidade

Além de sua formulação matemática robusta, o sucesso do XGBoost se deve a uma série de otimizações de sistema que garantem sua escalabilidade. Uma inovação chave é o seu algoritmo de busca de divisão ciente da esparsidade (*sparsity-aware*). Em muitos problemas do mundo real, os dados de entrada são esparsos. O XGBoost trata valores

ausentes de forma unificada, aprendendo uma direção padrão para cada nó da árvore, o que permite lidar com a esparsidade de forma eficiente sem a necessidade de pré-processamento ou imputação de dados (CHEN; GUESTRIN, 2016).

Para lidar com grandes conjuntos de dados que não cabem na memória, o XGBoost emprega uma estrutura de blocos para computação fora do núcleo (*out-of-core*). Os dados são divididos em blocos e armazenados em disco, sendo comprimidos e lidos de forma concorrente ao treinamento. Além disso, o algoritmo utiliza uma estrutura de dados de coluna pré-ordenada e otimizações cientes do cache (*cache-aware*) para minimizar os gargalos de acesso à memória e acelerar o cálculo das estatísticas de gradiente necessárias para encontrar as melhores divisões. Essas otimizações de sistema, combinadas com a capacidade de computação paralela e distribuída, permitem que o XGBoost treine modelos em conjuntos de dados de bilhões de exemplos de forma eficiente.

2.10 Redes Neurais

Rede Neural Artificial (RNA) são modelos computacionais inspirados na estrutura e no funcionamento do cérebro biológico, projetados para aprender a partir de dados. Elas são compostas por unidades de processamento interconectadas, chamadas de neurônios artificiais, que são organizadas em camadas. O objetivo fundamental de uma RNA é ajustar a força das conexões (os pesos) entre os neurônios para realizar tarefas complexas, como classificação, regressão e reconhecimento de padrões, aprendendo representações hierárquicas dos dados (GOODFELLOW; BENGIO; COURVILLE, 2016).

A profundidade da rede, ou seja, o número de camadas ocultas, é um fator crucial que distingue as Redes Neurais Profundas Deep Neural Network (DNN)s das arquiteturas mais rasas. Redes profundas permitem que o modelo aprenda uma hierarquia de características, onde cada camada constrói representações cada vez mais abstratas e complexas a partir das saídas da camada anterior. Essa capacidade de aprender características em múltiplos níveis de abstração é o que confere às DNNs seu poder para resolver problemas complexos em domínios como visão computacional e processamento de linguagem natural (LECUN; BENGIO; HINTON, 2015).

2.10.1 Arquitetura e Componentes Fundamentais

A unidade básica de uma RNA é o neurônio artificial. Cada neurônio recebe um ou mais sinais de entrada, realiza uma combinação linear desses sinais ponderada por seus respectivos pesos, adiciona um termo de viés (*bias*) e, em seguida, aplica uma função de ativação não linear para produzir um sinal de saída. Matematicamente, a saída a de um neurônio pode ser descrita como:

$$a = \sigma \left(\sum_{i=1}^n w_i x_i + b \right) = \sigma(w^T x + b) \quad (45)$$

onde x é o vetor de entradas, w é o vetor de pesos sinápticos, b é o viés e $\sigma(\cdot)$ é a função de ativação. Os pesos w_i modulam a influência de cada entrada x_i , enquanto o viés b atua como um limiar de ativação, deslocando a função de ativação. A organização desses neurônios em camadas — uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída — constitui a arquitetura da rede. Em uma rede *feedforward*, a informação propaga-se unidirecionalmente da entrada para a saída, sem ciclos (MAIER et al., 2019).

As funções de ativação são um componente crítico, pois introduzem não-linearidades no modelo, permitindo que a rede aprenda mapeamentos complexos que vão além das capacidades dos modelos lineares. Funções clássicas como a sigmoide e a tangente hiperbólica ($\tanh$) são limitadas por problemas de saturação, onde o gradiente se aproxima de zero para valores de entrada muito grandes ou muito pequenos, dificultando o treinamento de redes profundas (o problema do desaparecimento do gradiente). Para mitigar isso, a Unidade Linear Retificada (ReLU), definida como $f(x) = \max(0, x)$, tornou-se a escolha padrão na maioria das aplicações modernas. A ReLU é computacionalmente eficiente e não satura na região positiva, facilitando o fluxo de gradientes. Variantes como a Leaky ReLU e a Parametric ReLU (PReLU) foram propostas para resolver o problema do neurônio morto, permitindo um pequeno gradiente quando a unidade está inativa (GOODFELLOW; BENGIO; COURVILLE, 2016).

2.10.2 Aprendizado e Otimização

O processo de treinamento de uma RNA é um problema de otimização, onde o objetivo é encontrar o conjunto de parâmetros (pesos e vieses) que minimiza uma função de perda (ou custo). A função de perda, $L(\hat{y}, y)$, quantifica a discrepância entre a saída prevista pela rede, $\hat{y}$, e o valor real (alvo), y . A escolha da função de perda depende da tarefa: para regressão, o erro quadrático médio (MSE) é comum, enquanto para classificação, a entropia cruzada categórica é a mais utilizada, pois está diretamente relacionada à maximização da verossimilhança dos dados de treinamento (GOODFELLOW; BENGIO; COURVILLE, 2016).

O algoritmo de otimização mais utilizado para treinar redes neurais é o Gradiente Descendente Estocástico (SGD). Em vez de calcular o gradiente da função de perda sobre todo o conjunto de dados (o que seria computacionalmente proibitivo para grandes datasets), o SGD estima o gradiente utilizando um pequeno subconjunto de dados, chamado de *minibatch* (GOODFELLOW; BENGIO; COURVILLE, 2016). O gradiente da perda em relação a cada parâmetro da rede é calculado eficientemente pelo algoritmo de retropropagação (*backpropagation*), que aplica a regra da cadeia do cálculo de forma recursiva, da camada de saída para a de entrada. Os parâmetros são então atualizados na direção oposta ao gradiente, com a magnitude do passo controlada por um hiperparâmetro chamado taxa de aprendizado (η) (GOODFELLOW; BENGIO; COURVILLE, 2016).

Embora o SGD seja o pilar do treinamento de redes profundas, diversas variantes foram desenvolvidas para acelerar a convergência e melhorar a robustez. Algoritmos como o Momentum, RMSProp e Adam introduzem mecanismos adaptativos para a taxa de aprendizado, ajustando-a individualmente para cada parâmetro com base no histórico de seus gradientes (SAHINER et al., 2018). Esses otimizadores adaptativos frequentemente convergem mais rápido e são menos sensíveis à escolha inicial da taxa de aprendizado, tornando-se a opção padrão em muitas aplicações práticas (SAHINER et al., 2018).

2.10.3 Redes Neurais Siamesas

Redes Neurais Siamesas (*Siamese Neural Networks*) representam uma arquitetura especializada projetada para aprender uma função de similaridade ou distância entre duas entradas comparáveis (BROMLEY et al., 1993). Em vez de treinar uma rede para classificar uma entrada em uma de várias categorias, uma rede siamesa aprende a determinar se duas entradas são semelhantes ou diferentes.

A arquitetura consiste em duas redes gêmeas idênticas que compartilham exatamente os mesmos pesos e parâmetros. Cada rede recebe uma das duas entradas a serem comparadas e as processa de forma independente para gerar um vetor de características, ou *embedding*. Esses *embeddings* são vetores em um espaço de alta dimensão que representam as características salientes da entrada. O objetivo é que, nesse espaço de *embedding*, as representações de entradas semelhantes fiquem próximas, enquanto as representações de entradas diferentes fiquem distantes (ROY et al., 2019).

Formalmente, dadas duas entradas x_1 e x_2 , as duas sub-redes gêmeas, parametrizadas pela mesma função de *embedding* $f_W(x)$, produzem os vetores de características $h_1 = f_W(x_1)$ e $h_2 = f_W(x_2)$. A similaridade entre x_1 e x_2 é então calculada como uma função da distância entre h_1 e h_2 , como a distância Euclidiana.

O treinamento de redes siamesas é frequentemente realizado utilizando uma função de perda contrastiva (*contrastive loss*) ou uma perda de tripletos (*triplet loss*). A perda contrastiva incentiva a rede a produzir representações próximas para pares de entradas da mesma classe (pares positivos) e representações distantes para pares de classes diferentes (pares negativos) (HADSELL; CHOPRA; LECUN, 2006). A perda de tripletos, por sua vez, opera sobre um trio de exemplos: uma âncora (x_a), um exemplo positivo (x_p) da mesma classe que a âncora, e um exemplo negativo (x_n) de uma classe diferente. O objetivo é que a distância entre a âncora e o positivo seja menor do que a distância entre a âncora e o negativo, por uma margem de segurança τ :

$$\|f_W(x_a) - f_W(x_p)\|_2^2 + \tau < \|f_W(x_a) - f_W(x_n)\|_2^2 \quad (46)$$

As redes siamesas são particularmente úteis em cenários onde o número de classes é muito grande ou indefinido, e onde o objetivo principal é a verificação ou a recuperação

baseada em similaridade, como na verificação de assinaturas, reconhecimento facial e recuperação de imagens (KOCH, 2015). No contexto de imagens médicas, elas podem ser aplicadas para comparar imagens de um mesmo paciente ao longo do tempo, ou para comparar um paciente com um banco de dados de casos conhecidos, aprendendo uma métrica de similaridade que captura mudanças sutis ou padrões patológicos.

2.11 Capacidade, *Overfitting* e Regularização

A capacidade de um modelo de aprendizado de máquina refere-se à sua habilidade de se ajustar a uma ampla variedade de funções. Modelos com alta capacidade podem modelar relações complexas, mas correm o risco de *overfitting* (sobreajuste), que ocorre quando o modelo aprende o ruído e as particularidades do conjunto de treinamento em vez do padrão subjacente, resultando em um baixo desempenho em dados não vistos. Por outro lado, modelos com baixa capacidade podem sofrer de *underfitting* (subajuste), não conseguindo capturar a complexidade dos dados e apresentando um desempenho ruim tanto no treinamento quanto no teste (GOODFELLOW; BENGIO; COURVILLE, 2016).

O desafio central no treinamento de modelos de aprendizado de máquinas é encontrar um equilíbrio que minimize o erro de generalização, ou seja, o erro em dados novos. A regularização compreende um conjunto de técnicas projetadas para reduzir o erro de generalização, muitas vezes à custa de um aumento no erro de treinamento. Essas técnicas impõem restrições ao algoritmo de aprendizado, expressando uma preferência por modelos mais simples e, portanto, mais propensos a generalizar (GOODFELLOW; BENGIO; COURVILLE, 2016).

As técnicas de regularização mais comuns incluem a penalização de normas dos pesos, como a regularização L2 (decaimento de peso ou *weight decay*), que penaliza pesos grandes e incentiva a rede a encontrar soluções mais suaves, e a regularização L1, que promove a esparsidade nos pesos (fazendo com que muitos se tornem zero). Outra técnica poderosa, especialmente para redes neurais, é o *Dropout*, que, durante o treinamento, desativa aleatoriamente um subconjunto de neurônios em cada iteração. Isso força a rede a aprender características redundantes e robustas, impedindo que os neurônios co-adaptem-se excessivamente aos dados de treinamento (SRIVASTAVA et al., 2014). A parada antecipada (*early stopping*), onde o treinamento é interrompido quando o desempenho em um conjunto de validação começa a piorar, é outra forma eficaz e computacionalmente barata de regularização.

2.12 Análise estatística

A análise estatística tem como objetivo permitir uma melhor interpretação dos dados e na validação de hipóteses em projetos de pesquisa. Em estudos com dados categóricos,

testes estatísticos específicos são empregados para avaliar a associação entre variáveis ou comparar proporções.

2.12.1 Teste de McNemar

O teste de McNemar (MCNEMAR, 1947) é um teste estatístico não-paramétrico especificamente projetado para analisar dados categóricos pareados ou relacionados, com um resultado binário. Sua principal aplicação em aprendizado de máquina, especialmente em tarefas de classificação binária, é comparar o desempenho de dois modelos quando aplicados ao mesmo conjunto de dados (MCNEMAR, 1947; EDWARDS, 1948).

A característica distintiva do teste de McNemar é que ele se concentra nas instâncias em que os dois modelos (ou classificadores) discordam. Em vez de uma avaliação geral da acurácia, ele investiga se a natureza dos erros cometidos por um modelo difere significativamente da do outro (MCNEMAR, 1947; EDWARDS, 1948).

A hipótese nula (H_0) do teste de McNemar estabelece que não há diferença estatisticamente significativa nas taxas de erro dos dois modelos. Mais precisamente, ele postula que a proporção de instâncias onde o Modelo A está correto e o Modelo B está incorreto é igual à proporção de instâncias onde o Modelo B está correto e o Modelo A está incorreto. A hipótese alternativa (H_1) sugere que existe uma diferença significativa nessas proporções (MCNEMAR, 1947; EDWARDS, 1948).

O teste é construído a partir de uma tabela de contingência 2×2 que sumariza as previsões de dois modelos em relação aos rótulos verdadeiros. No contexto de comparação de classificadores, essa tabela é popularmente conhecida como Tabela de McNemar (EDWARDS, 1948). A estrutura da tabela é a seguinte:

Tabela 1 – Tabela de Contingência para o Teste de McNemar

	Modelo B Correto	Modelo B Incorreto
Modelo A Correto	a	b
Modelo A Incorreto	c	d

Onde:

- a : Número de instâncias em que ambos os modelos classificaram corretamente.
- b : Número de instâncias em que o Modelo 1 classificou corretamente, mas o Modelo 2 classificou incorretamente.
- c : Número de instâncias em que o Modelo 1 classificou incorretamente, mas o Modelo 2 classificou corretamente.
- d : Número de instâncias em que ambos os modelos classificaram incorretamente.

O estatístico de teste de McNemar (χ^2) é calculado com base nos valores das células discordantes (b e c) da tabela de contingência. A fórmula utilizada, que inclui uma correção de continuidade para melhor aproximação da distribuição qui-quadrado, é:

$$\chi^2 = \frac{(|b - c| - 1)^2}{b + c} \quad (47)$$

O valor de χ^2 é então comparado com uma distribuição qui-quadrado com 1 grau de liberdade para obter um p -valor. Um p -valor menor ou igual ao nível de significância predefinido (usualmente $\alpha = 0.05$) leva à rejeição da hipótese nula. A interpretação da significância e do modelo superior é baseada nessa comparação e na análise de b vs. c (MCNEMAR, 1947; EDWARDS, 1948).

A importância do teste de McNemar em aprendizado de máquina reside em sua capacidade de ir além da acurácia geral, permitindo uma análise mais aprofundada dos tipos de erros e das diferenças no comportamento dos modelos em dados pareados. A aplicação desse teste garante que a análise seja feita sobre o conjunto completo de predições, considerando o pareamento dos dados.

Capítulo 3

Revisão Bibliográfica

3.1 Assimetrias inter-hemisféricas

Acredita-se que assimetrias estruturais e funcionais são características benéficas para o cérebro, amplamente aceitas e confirmadas em estudos por autópsia e imagens *in vivo* (LONG et al., 2013). Alterações da assimetria hemisférica também ocorrem com o envelhecimento saudável, lesões cerebrais ou distúrbios neurodegenerativos patológicos nos quais a integridade anatômica e funcional do cérebro é prejudicada (LONG et al., 2013; Alzheimer's Association, 2023; NEMATİ et al., 2023).

Nesse contexto, vários autores buscam realizar a análise de tais assimetrias para desenvolver biomarcadores de doenças neurodegenerativas como o Alzheimer (WACHINGER et al., 2015; WACHINGER et al., 2016; LONG; JIANG; ZHANG, 2018; OLIVEIRA; POLONI; FERRARI, 2020; POLONI et al., 2021; KONG et al., 2022). As técnicas existentes para a análise da neurodegeneração na DA utilizam abordagens diversas, indo desde o uso de informações volumétricas em determinadas regiões (KONG et al., 2018; KONG et al., 2022; LONG; JIANG; ZHANG, 2018) até a criação de modelos com técnicas de aprendizado de máquina para o processamento da imagem de RM (OLIVEIRA; POLONI; FERRARI, 2020; POLONI et al., 2021).

No contexto desta pesquisa, os trabalhos apresentados na próxima seção foram selecionados pois são estudos voltados à análise de assimetrias inter-hemisféricas em diferentes populações, incluindo a DA.

3.1.1 Trabalhos correlatos

A associação Enhancing NeuroImaging Genetics through Meta-Analysis (ENIGMA), Kong et al. (2018), estudou imagens de 99 bases de dados totalizando 17.141 indivíduos saudáveis para analisar as assimetrias no córtex cerebral. Esse estudo avaliou a espessura e a área superficial do córtex em imagens de RM, revelando padrões significativos de assimetria tanto em nível hemisférico quanto regional. Notadamente, foi observado um córtex geralmente mais espesso, mas com menor área superficial no hemisfério esquerdo em comparação com o direito. O processamento dos dados foi realizado em imagens de RM T1-w com o *pipeline recon-all*¹ implementado no software FreeSurfer², as etapas principais do pré-processamento consiste na normalização de intensidades, extração do encéfalo, segmentação das substâncias branca e cinzentas e delineação de suas superfícies. Em seguida, essas superfícies são separadas de acordo com as diferentes regiões do córtex via rotulação automática baseada num atlas de probabilidade, curvaturas locais e informações contextuais das geometrias específicas de cada indivíduo. Com isso, são calculadas a área de cada superfície e espessura do córtex em 68 regiões (34 para cada hemisfério), utilizando o referencial espacial de cada indivíduo. Para calcular o índice de assimetria (IA), utilizou-se a fórmula

$$IA = \frac{L - R}{\left(\frac{L+R}{2}\right)}, \quad (48)$$

em que L e R são os valores correspondentes de espessura ou área dos hemisférios esquerdo e direito, respectivamente. Assim, valores positivos ou negativos revelam informações sobre a direção da lateralização hemisférica para esquerda e direita, respectivamente. Adicionalmente, também foram realizados testes estatísticos *t-Student* e *d* de Cohen para medir diferenças inter-hemisféricas e estimar o efeito da assimetria no nível da população. Os resultados obtidos para cada meta-análise são sumarizados a seguir:

❑ **Assimetria no nível populacional:** A análise revelou assimetrias generalizadas em espessura cortical e área superficial. Em geral, o córtex é mais espesso no hemisfério esquerdo mas apresenta uma área superficial menor em comparação com o hemisfério direito. Para a espessura do córtex, 76,5% das regiões mostraram alterações significativas na assimetria da espessura, com lateralização para a esquerda (i.e., esquerda > direita) nas regiões do córtex anterior, incluindo o córtex pré-frontal medial. A lateralização para a direita foi observada nas regiões do córtex posterior, incluindo o lóbulo parietal e occipital. Os autores destacam que essa forte assimetria frontal-posterior pode estar relacionada com a rotação de Yakovlevian (KUO; MASSOUD, 2022).

❑ **Efeitos de gênero na espessura do córtex:** Não foi observado efeito global significativo do sexo no IA para a espessura cortical, mas efeitos de regiões espe-

¹ <<https://surfer.nmr.mgh.harvard.edu/fswiki/recon-all>>

² <<https://surfer.nmr.mgh.harvard.edu/>>

cíficas foram encontrados nas regiões do giro parahipocampal e córtex entorrinal, indicando maior lateralização para a esquerda no sexo masculino. Em contraste, a área superficial apresentou diferença significativa na assimetria global, indicando que os homens têm, em média, mais assimetria para a direita em comparação com as mulheres. Além disso, esses efeitos de assimetria foram agravados em amostras mais jovens em comparação com as de idade avançada, sendo o sexo feminino o mais afetado.

- ❑ **Idade e sua relação com as assimetrias no córtex:** Inicialmente, não foram observados efeitos significativos da idade nas assimetrias globais. Entretanto, em conjuntos de dados com faixas etárias mais amplas, o aumento da idade foi associado com uma assimetria esquerda mais pronunciada na espessura cortical no giro temporal superior e córtex entorrinal.
- ❑ **Diferenças de grupo por preferência motora (mão dominante):** O estudo não encontrou associações significativas entre preferência motora e assimetrias corticais.
- ❑ **Efeitos do volume intracraniano na assimetria do córtex:** O volume intracraniano mostrou um efeito positivo significativo na assimetria geral da espessura cortical, aumento da lateralização para a esquerda. Efeitos específicos foram encontrados para as regiões do lóbulo parietal inferior e ínsula.
- ❑ **Hereditariedade das assimetrias anatômicas corticais:** As assimetrias hemisféricas gerais de espessura cortical e área superficial mostraram baixa hereditariedade. As assimetrias mais hereditárias na espessura cortical regional foram encontradas no istmo e no giro cingulado anterior.

A meta-análise de Kong et al. (2022) é uma extensão do estudo anterior de Kong et al. (2018), aprofundando-se na análise das assimetrias cerebrais em condições de saúde e doenças. O estudo utilizou imagens de RM T1-w obtidas nos últimos 20 anos e inclusas no consórcio ENIGMA, e aplicou a metodologia *recon-all* mencionada anteriormente para as etapas de pré-processamento, extrações das 34 regiões de interesse de cada hemisfério, e cálculo das espessuras e área de superfície das regiões corticais. Nesse estudo, os autores realizaram uma modificação no IA proposto por Kong et al. (2018), que foi a substituição da média dos valores de L e R do denominador pela soma simples entre eles, resultando na equação $AI = (L - R)/(L + R)$.

Uma determinada região é considerada assimétrica em nível populacional quando a média do IA é levemente maior que zero ou, em alguns estudos da meta-análise, quando o teste t de *Student* apresenta diferenças significativas. Os resultados desse estudo confirmaram as análises de Kong et al. (2018) para a população cognitivamente normal, obtendo os resultados para a população geral quanto à lateralização para a esquerda na espessura do córtex e lateralização para a direita na área superficial, e resultados similares

quando comparados com os outros fatores de sexo, idade e preferência manual. Ademais, esse estudo comparou os resultados de cérebros cognitivamente normais com cérebros de indivíduos neurodivergentes em três diferentes populações: (a) transtorno do espectro autista (TEA); (b) transtorno depressivo maior (TDM); (c) transtorno obsessivo-compulsivo (TOC). Na população TEA, os resultados indicam redução nas assimetrias, com efeitos pequenos e variáveis, não sugerindo seu uso como biomarcador. Para a população TDM, a análise de grande escala não encontrou diferenças significativas na assimetria estrutural do cérebro. Nos casos de TOC, mudanças na assimetria foram observadas especialmente em pacientes pediátricos, indicando alterações laterais no neurodesenvolvimento dos circuitos que envolvem a região do tálamo.

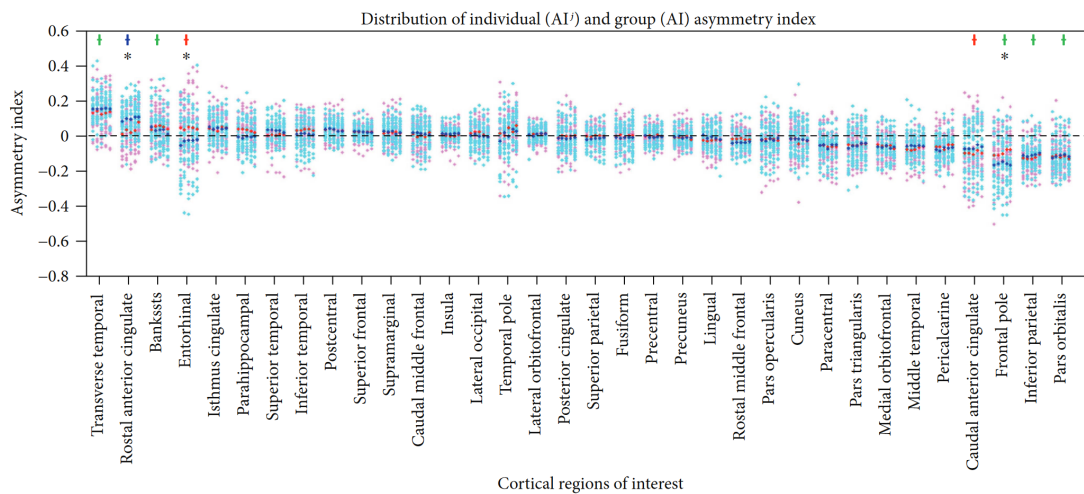
O CCL é uma etapa intermediária entre o declínio cognitivo do envelhecimento natural e aquele causado pela DA (LONG; JIANG; ZHANG, 2018). Pacientes com CCL possuem chances entre três à cinco vezes maiores de desenvolverem a DA, no entanto, nem todas as pessoas diagnosticadas com CCL evoluem para a DA (denominados CCL conversores), sendo possível a recuperação e reversão desse estágio (LONG; JIANG; ZHANG, 2018), e identificar os casos CCL conversores é fundamental para entender o progresso da doença e iniciar os tratamentos preventivos. Dessa forma, Long, Jiang e Zhang (2018) utilizam biomarcadores morfológicos extraídos da assimetria em dados longitudinais de RM em 53 pacientes com CCL conversores, tendo como objetivo criar um índice de assimetria para capaz de diferenciar entre esses dois grupos.

O método de Long, Jiang e Zhang (2018) foi testado em um conjunto de imagens 3D de RM ponderadas em T1-w obtidas da base de dados ADNI, retratando 26 pacientes CCL conversores e 27 CCL estáveis, com idades de 73.35 ± 6.71 anos. Todos os pacientes eram destros e passaram por pelo menos cinco visitas subsequentes no intervalo de aproximadamente um ano. O grupo CCL estáveis não foi diagnosticado com DA em nenhuma das visitas. O grupo CCL conversores permaneceu com diagnóstico de CCL durante pelo menos três visitas consecutivas antes do diagnóstico de Alzheimer, que foi mantido por duas ou mais visitas seguintes após a conversão. As imagens passaram pelo pré-processamento para normalização de intensidade extração do encéfalo e alinhamento espacial. Ainda, para cada imagem, o software *FreeSurfer* foi utilizado para extrair 34 regiões de interesse. Para cada região de interesse (*region of interest* - ROI), em cada hemisfério, foram obtidos a área de superfície, o índice médio de curvatura, e a espessura do córtex. Essas medidas foram usadas para o teste de diferenças hemisféricas entre os grupos via análise de covariância multivariada. O índice de assimetria de cada paciente, para cada região, é calculado como

$$AF^j = \sum_i \frac{L_i^j - R_i^j}{L_i^j + R_i^j}, \quad (49)$$

sendo que a média dessa fórmula é utilizada para identificar a variação em cada grupo (LONG; JIANG; ZHANG, 2018). Nesse caso, um índice positivo indica lateralização para

Figura 6 – Índice de assimetria em diferentes estruturas anatômicas.



Fonte: Long, Jiang e Zhang (2018).

esquerda enquanto que valores negativos indicam para a direita. A Figura 3.1.1 apresenta a distribuição desse índice para diferentes regiões anatômicas.

Diferentes padrões de assimetria foram identificados em regiões do sistema límbico, incluindo o córtex cingulado anterior e o córtex entorrinal. O grupo CCL estável apresentou, em todas as visitas, lateralização para a esquerda no córtex entorrinal e giro cingulado anterior. Nesse mesmo grupo foi detectada lateralização para a direita no giro cingulado anterior caudal no início, na segunda e quarta visitas. Os resultados indicam uma redução no índice de assimetria entre grupos em certas regiões, com tendência a zero ao longo do tempo. Essas regiões incluem áreas relacionadas à linguagem e memória, como o sulco temporal superior e córtex entorrinal, respectivamente. Os casos CCL conversores exibiram uma tendência decrescente mais notável do que os estáveis, especialmente em áreas das margens do sulco temporal superior, o córtex entorrinal e o parte triangular. Análises de teste t pareado demonstraram que diferenças significativas entre grupos foram encontradas no córtex entorrinal, polo frontal e cíngulo anterior rostral. Em suma, de acordo com Long, Jiang e Zhang (2018), a assimetria no córtex entorrinal pode ser utilizada como um biomarcador significativo para diferenciar entre CCL conversor e estável.

Em Wachinger et al. (2015), os autores propuseram um novo método para analisar o formato de estruturas cerebrais utilizando uma técnica denominada técnica denominada *BrainPrint* Reuter e Fischl (2011). Diferentemente de análises volumétricas, a técnica proposta seria mais robusta a variações de intensidades ocasionadas por ruídos e mais sensível a variações sutis inter-hemisféricas. A técnica consiste em segmentar as estruturas anatômicas do cérebro com o software *FreeSurfer* e criar malhas tridimensionais das estruturas corticais e subcorticais de interesse para servirem como entrada para o descritor *shapeDNA*. O descritor busca extrair uma representação compacta e discriminativa da morfologia cerebral, capturando informações de forma de um conjunto de estruturas corticais e subcorticais. Ele se baseia na solução do problema de autovalores do ope-

radar Laplace-Beltrami nas malhas do *FreeSurfer*. Uma característica importante do *shapeDNA* é que pequenas mudanças na forma do objeto resultam em pequenas variações no descritor, tornando-o um método robusto para comparação de formas. Para análise de assimetrias, foram utilizadas as distâncias entre descritores extraídos de malhas triangulares de 12 estruturas do encéfalo da base ADNI. A métrica de distância é calculada como $d_i = \|\lambda_i^{\text{esq}} - \lambda_i^{\text{dir}}\|_2$, sendo λ_i um dos 50 autovalores em cada hemisfério. Devido sua propriedade isométrica, a técnica *shapeDNA* permite a comparação direta das sequências de autovalores, já que objetos espelhados geram o mesmo *shapeDNA*. Na técnica proposta, primeiro os autores calcularam o índice médio de assimetria de cada indivíduo da base e, em seguida, compararam as diferenças com todas as imagens da base. Ao analisar as diferenças entre estruturas corticais e subcorticais, observou-se que as estruturas corticais exibiam os menores rankings, indicando maior variabilidade e, portanto, menor assimetria em comparação com a similaridade intra-indivíduos. Estruturas subcorticais por outro lado, apresentaram um rank médio maior que 10%, isto é, dada uma subestrutura qualquer, a distância entre descritores em 10% dos indivíduos da base de dados é menor no hemisfério direito, indicando assimetria. Também foi constatado que 18,5% dos autovalores apresentavam assimetria direcional significativa. Ao analisarem a influência de idade e sexo, os autores observaram assimetrias significantes em algumas estruturas, com tendência de aumento assimétrico nas idades mais avançadas, e está presente mais fortemente nos homens.

Wachinger et al. (2016) utilizam a metodologia do *BrainPrint* (WACHINGER et al., 2015) para realizar uma análise mais aprofundada do impacto da assimetria na DA. A pesquisa empregou um índice de assimetria com os autovalores de cada hemisfério dada por $Y_s = \|\bar{\lambda}_s^{\text{esq}} - \bar{\lambda}_s^{\text{dir}}\|_{\Sigma_s}$, sendo Σ_s a matriz de covariância para a estrutura s em todos os indivíduos. Para a análise de conversão longitudinal foram selecionadas imagens de RM da base ADNI, totalizando 697 indivíduos que possuem três ou mais imagens de diferentes sessões. Em seguida, um modelo linear de efeitos mistos, que incluiu a idade na primeira sessão de RM, a diferença em anos até a próxima sessão e o diagnóstico como variável categórica (CN, CCL não-conversores, CCL conversores e DA), foi utilizado para estudar o efeito transversal e longitudinal na assimetria cerebral. Os resultados demonstram que a assimetria de forma é mais fortemente associada à idade, quantidade de anos desde a primeira sessão e diagnóstico em comparação com a assimetria de volume para a maioria das estruturas. Além disso, a assimetria de do *BrainPrint* nas estruturas do hipocampo e amígdala difere significativamente em todos os grupos diagnósticos, sugerindo um aumento da assimetria com o avanço da doença. O estudo também explorou a associação entre o aumento da assimetria e as mudanças na espessura cortical, utilizando um modelo de efeitos mistos ligeiramente diferente, e confirma que a atrofia da substância cinzenta está significativamente associada a aumentos intra-individuais na assimetria do hipocampo. A assimetria intra-individual no grupo de controle (CN) aumentou apenas

em relação significativa com idade. Além disso, utilizando o modelo de riscos proporcionais de Cox (COX, 1972), a pesquisa avaliou a progressão do CCL para a DA em 374 indivíduos com CCL. As estruturas que apresentaram resultados significativos para o descritor *BrainPrint* foram: estruturas corticais, hipocampo, amígdala e caudato.

Oliveira, Poloni e Ferrari (2020) introduziram uma nova abordagem para a análise de assimetrias na região hipocampal utilizando bancos de filtros log-Gabor 3D. Na metodologia proposta, os autores utilizaram uma máscara binária dilatada na base NAC³, alinhada às imagens de análise da base ADNI via correção de registro, para identificar a região hipocampal nas imagens. Em seguida, um banco de filtros log-Gabor 3D é aplicado para identificar as respostas direcionais em diferentes escalas na região hipocampal. Para cada orientação da imagem filtrada, quatro momentos estatísticos (média, variância, *skewness* e *kurtosis*) são extraídos das distribuições de intensidade na região do hemisfério esquerdo rotacionado e direito. Duas abordagens foram empregadas para construir um vetor de características que contenha as informações a respeito da assimetria inter-hemisférica: (i) diferença absoluta entre os elementos dos vetores, (ii) multiplicação entre os elementos dos vetores. Na primeira abordagem, a diferença absoluta entre o vetor de características do hemisfério esquerdo $x_l^{f,s}$ e do hemisfério direito $x_r^{f,s}$, é calculada como

$$v^{f,s}[k] = |x_l^{f,s}[k] - x_r^{f,s}[k]|, \quad (50)$$

em que $k = 0, \dots, N_0 - 1$ é o índice para cada elemento dos vetores, e $s = 0, \dots, N_{s-1}$ são as escalas da imagem filtrada. Assim, obtém-se o vetor de características em todas as escalas denotado por $y^f = [v^{f,0}; \dots, v^{f,N_{s-1}}]$ (OLIVEIRA; POLONI; FERRARI, 2020). Uma representação visual do vetor resultando pode ser visualizada na Figura 7.

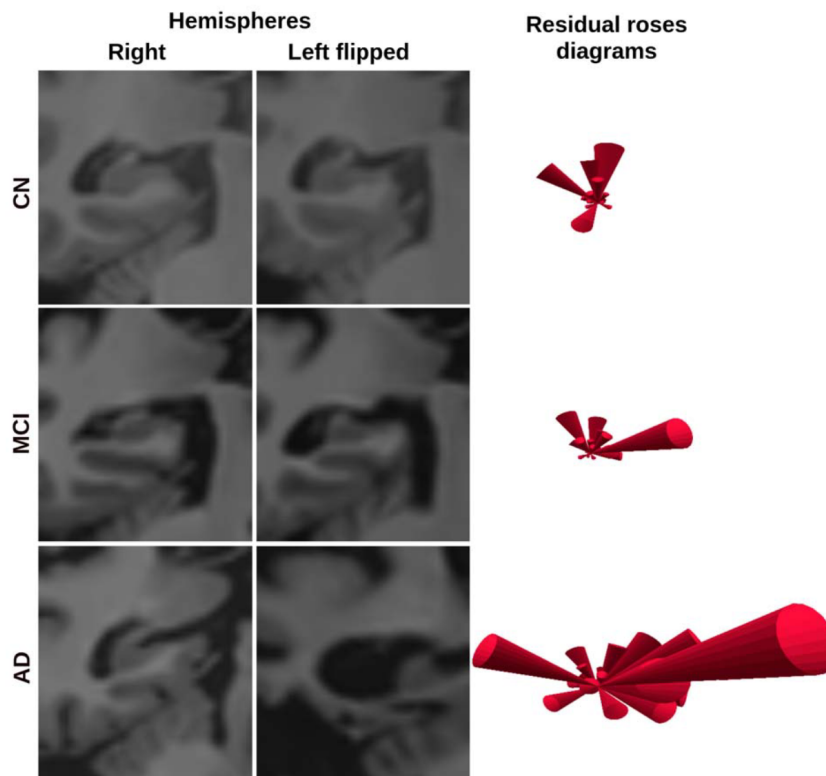
A segunda abordagem desenvolvida por Oliveira, Poloni e Ferrari (2020) foi baseada no estudo de Lee et al. (2014) que demonstrou que a multiplicação entre vetores de atributos é capaz de auxiliar tarefas de correspondências (casamentos) em técnicas de aprendizado de máquina. Neste contexto, Oliveira, Poloni e Ferrari (2020) definem o vetor de características contendo as informações de assimetria como

$$v^{f,s}[k] = x_l^{f,s}[k] \cdot x_r^{f,s}[k], \quad (51)$$

sendo o vetor resultante para todas as escalas s representado por $z^f = [v^{f,0}; \dots, v^{f,N_{s-1}}]$. Os vetores resultantes y^f e z^f foram utilizados para treinar e testar classificadores SVM na tarefa de diferenciar entre os grupos diagnósticos CN vs. DA e CN vs. CCL. Os dados utilizados para a obtenção desses vetores foram 762 imagens T1-w de RM da base ADNI, pertencentes à indivíduos entre os setenta e oitenta e cinco anos para os três grupos diagnósticos. Classificadores que utilizaram os vetores y^f obtiveram uma acurácia máxima de 64,19% na classificação dos grupos CN vs. AD, os resultados obtidos para a classificação entre os grupos CN vs. CCL não foram apresentados devido à baixa

³ Disponível em: <https://nac.spl.harvard.edu> Acesso: 01 dez, 2023.

Figura 7 – Diagrama de rosas da diferença absoluta entre os vetores de características do hemisfério esquerdo e direito.



Fonte: Oliveira, Poloni e Ferrari (2020).

eficácia segundo os autores. Os classificadores que utilizaram os vetores z^f obtiveram uma acurácia máxima de $71,23\% \pm 7,08\%$ e $81,43\% \pm 4,86\%$ para as tarefas de classificação entre os grupos CN vs. CCL e CN vs. DA, respectivamente. Esses resultados mostram que atributos indicativos de assimetria inter-hemisférica na região hipocampal apresentam o potencial para tarefas de classificação da DA em imagens de RM (OLIVEIRA; POLONI; FERRARI, 2020).

As técnicas desenvolvidas por Oliveira, Poloni e Ferrari (2020) foram ampliadas no estudo Poloni et al. (2021) para a classificação automática da DA em imagens de RM, utilizando apenas os atributos que apresentaram diferenças significativas entre os grupos de diagnóstico CN, CCL e DA no teste de análise de variância (*Analysis of Variance* -) (MISHRA et al., 2019). O método utilizado por Poloni et al. (2021) realiza o pré-processamento das imagens volumétricas de RM, seguida pela extração dos dois hemisférios cerebrais e a inversão do hemisfério esquerdo para que ambos tenham aproximadamente a mesma localização espacial. Após isso, na etapa de filtragem direcional, é realizada uma convolução dos hemisférios da imagem clínica com um banco de filtros log-Gabor 3D e todas as respostas direcionais são salvas. Em seguida, região hipocampal é extraída via multiplicação com uma máscara binária da base NAC⁴ alinhada às imagens

⁴ Disponível em: <https://nac.spl.harvard.edu/> Acesso: 01 dez, 2023.

clínicas via corregristo afim. As máscaras direita e esquerda são combinadas para levar em conta as variações espaciais das regiões e garantir o envolvimento de todas as estruturas hipocampais na análise de assimetria. Para isso, um operador de união é aplicado voxel a voxel às máscaras dos hemisférios, e a máscara resultante é expandida com uma operação de dilatação morfológica, usando uma esfera 3D de dois voxels de raio como elemento estruturante. Após a etapa de filtragem direcional, os primeiros quatro momentos estatísticos (média, variância, *skewness* e *kurtosis*) foram computados a partir dos valores de intensidade dentro da máscara hipocampal. Para medir a similaridade entre os atributos extraídos de cada hemisfério, a correlação cruzada normalizada (*Normalized Cross Correlation* - NCC) é calculada entre os vetores de características do hemisfério esquerdo ($\mathbf{a}_L$) e direito rotacionado ($\mathbf{a}_{R^*}$), ambos de tamanho $N = 26$. Essa operação de similaridade via NCC é definida como

$$NCC(x) = \frac{\sum_{i=0}^{N-1} \mathbf{a}_L[i] \cdot \mathbf{a}_{R^*}[i+x]}{\sqrt{\sum_{i=0}^{N-1} \mathbf{a}_L[i]^2} \sqrt{\sum_{i=0}^{N-1} \mathbf{a}_{R^*}[i+x]^2}}, \quad (52)$$

em que $x = 0, 1, \dots, N-1$ (POLONI et al., 2021). Essa operação ajuda a identificar a natureza da associação entre os dois hemisfério e como eles estão correlacionados entre si, avaliando não apenas a resposta direcional isoladamente, mas sim o padrão geral de assimetria que cada sujeito pode ter.

Na análise estatística, os autores propõem um novo índice de assimetria definido como

$$AI = 1 - \frac{1}{N} \sum_{i=0}^{N-1} NCC(x), \quad (53)$$

variando no intervalo $[0, 1]$, sendo 0 quando os vetores de atributos são idênticos e 1 quando não há nenhuma correlação entre os vetores de atributos (total assimetria). O teste estatístico *One-way* ANOVA é utilizado para avaliar se há diferença estatística significativa do IA entre os grupos de diagnóstico CN, CCL e DA para os diferentes atributos extraídos. Nos casos em que a hipótese nula é rejeitada (há diferenças em pelo menos um dos grupos), o teste de significância Turkey's HSD (ABDI; WILLIAMS, 2010) é utilizado para identificar quais os grupos que diferem entre si.

Para avaliar essa metodologia, Poloni et al. (2021) utilizaram imagens de RM da base ADNI de indivíduos de 60 à 85 anos, pertencentes aos grupos diagnósticos CN, CCL e DA, totalizando 762 imagens. Os atributos que apresentaram significância estatística foram utilizados para treinar e testar classificadores SVM à identificar entre as classes CN×CCL e CN×DA. Os melhores resultados obtidos para a classificação entre CN e CCL foi a acurácia de $69,44\% \pm 4,58\%$, e $82,59\% \pm 6,49\%$ para a classificação entre CN e DA. Os autores também realizaram uma comparação do classificador com as métricas propostas com técnicas tradicionais de segmentação e cálculo volumétrico comumente utilizadas na literatura. Assim, o classificador obteve acurácia de $82,61\% \pm 6,48\%$ e $72,68\% \pm 4,08\%$ para a identificação entre as classes CN vs. AD e CN vs. CCL, respectivamente. Embora

os resultados de acurácia da metodologia proposta não sejam superiores àqueles obtidos com os atributos volumétricos, é necessário considerar que a análise volumétrica utilizou todas as ROI do cérebro, enquanto a metodologia proposta utilizou apenas os atributos de assimetria da região hipocampal, demonstrando ser uma métrica importante para o melhor entendimento da DA.

Korbmacher et al. (2024) investigaram as diferenças na assimetria cerebral relacionadas à idade em uma população de 48.040 indivíduos da base UK Biobank, utilizando dados de RM estrutural e de difusão. Os autores propõem um novo conceito de idade cerebral hemisférica (*hemispheric brain age* - HBA), uma abordagem na qual a idade cerebral é estimada de forma independente para cada hemisfério utilizando métricas de RM multimodais. A metodologia foi proposta como um marcador auxiliar à predição convencional de idade cerebral, permitindo uma avaliação da assimetria ao comparar diretamente as estimativas de HBA esquerda e direita. Com isso, o objetivo dos autores foi determinar como as assimetrias estruturais do cérebro se manifestam da meia-idade à velhice e de que forma a HBA poderia capturar essas alterações.

Os resultados indicaram que a maioria das métricas de GM e WM apresentou assimetria, e que essa assimetria tendeu a ser maior em idades mais avançadas. Por outro lado, a discrepância entre a HBA esquerda e direita apresentou uma relação inversa, sendo menor em indivíduos mais velhos. As predições de HBA mostraram-se semelhantes com as da idade cerebral convencional em termos de acurácia, validando sua robustez. Além disso, a pesquisa revelou diferenças significativas específicas ao sexo, com análises estratificadas mostrando efeitos hemisféricos de modalidade de RM opostos entre homens e mulheres. Esses resultados apontam para o valor de HBA como uma ferramenta auxiliar para avaliar o envelhecimento com base nas assimetrias inter-hemisféricas do encéfalo. Os autores ressaltam que as associações mais fortes da assimetria com a idade foram encontradas em estruturas subcorticais e límbicas frequentemente impactadas por doenças neurodegenerativas. Dessa forma, uma grande discrepância entre as HBAs de um mesmo indivíduo poderia, portanto, atuar como um biomarcador de desequilíbrio na saúde cerebral, com potencial de indicar possíveis patologias.

Em uma revisão sistemática sobre as implicações clínicas das assimetrias cerebrais, Ocklenburg et al. (2024) investigaram como as diferenças funcionais e estruturais entre os hemisférios cerebrais podem ser interpretadas no contexto clínicos. Os autores destacam que, embora as assimetrias sejam um princípio fundamental da organização cerebral, existe uma significativa variabilidade interindividual que precisa ser considerada. O estudo aborda três formas principais de assimetria: (i) funcional, como lateralização da linguagem; (ii) estrutural, como volumetria de regiões anatômicas; e (iii) conectiva, como a conectividade funcional e estrutural, que podem ser quantificadas via técnicas de neuroimagem. Essa variabilidade natural, segundo o estudo, representa um desafio significativo para a aplicação clínica direta desses marcadores.

A principal conclusão do estudo de Ocklenburg et al. (2024) é que as alterações na assimetria cerebral possuem valor limitado como biomarcadores de diagnóstico para transtornos neuropsiquiátricos e de neurodesenvolvimento, como esquizofrenia ou autismo. A falta de especificidade — dado que alterações similares ocorrem em diferentes condições e também em indivíduos saudáveis — impede seu uso como uma ferramenta de diagnóstico diferencial robusta. Por outro lado, os autores argumentam que as assimetrias cerebrais são mais promissoras como biomarcadores de sintomas e de resposta ao tratamento. Por exemplo, a lateralidade do início dos sintomas em doenças neurodegenerativas ou o padrão de assimetria funcional em um paciente podem prever manifestações clínicas específicas e orientar a aplicação de terapias lateralizadas, como a estimulação cerebral não invasiva e neurocirurgias.

Os resultados de Ocklenburg et al. (2024) reforçam a necessidade de desenvolver modelos computacionais que superem a simples classificação do caso-controle e se concentrem na predição de patologias clínicas individualizados. Nesse sentido, os autores recomendam mapear fenótipos de assimetria multifatoriais à gravidade dos sintomas e à reposta terapêutica, em vez de apenas ao diagnóstico. Para a área de visão computacional, isso implica que a extração de características de neuroimagens não deve visar apenas a identificação de padrões de grupo, mas sim a quantificação da variabilidade individual das assimetrias cerebrais, fornecendo uma base referencial para a possível criação de sistemas diagnósticos de apoio à decisão clínica.

Em um estudo transversal com 1.047 idosos residentes na comunidade, Matsuda et al. (2025) investigaram a associação entre a assimetria estrutural do cérebro e o comprometimento cognitivo. Foram utilizadas imagens de RM para avaliar os volumes de regiões cerebrais específicas, incluindo o hipocampo, a amígdala, o córtex pré-frontal dorsolateral e o córtex cingulado anterior. A análise da assimetria estrutural foi realizada através de um índice calculado a partir dos volumes dos hemisférios direito e esquerdo de cada região de interesse. A metodologia empregada envolveu o processamento das imagens cerebrais com o software FreeSurfer para a segmentação e cálculo dos volumes das estruturas de interesse. Para avaliar a relação entre a assimetria cerebral e o comprometimento cognitivo, os autores utilizaram a análise de regressão logística binomial. As funções cognitivas, como memória, atenção e função executiva, foram avaliadas por meio de testes neuropsicológicos, sendo o comprometimento definido como um desempenho 1,5 desvio padrão abaixo da média normativa.

Os autores destacam que, no conjunto de idosos com funções de memória relativamente preservadas e sem diagnóstico de demência, o comprometimento de memória não apresentou associação significativa com a assimetria estrutural cerebral. Em vez disso, a perda de memória esteve associada à redução do volume total do hipocampo e da amígdala. Estes resultados divergem de estudos anteriores que encontraram uma relação entre a assimetria do hipocampo e da amígdala e a gravidade da demência. A diferença nas características

dos participantes do estudo é a principal hipótese para justificar essa discrepância nos resultados. Assim, para a população estudada, o volume total dessas estruturas, e não a sua assimetria, mostrou-se um indicador mais relevante para o comprometimento da memória. Por outro lado, o estudo revelou uma associação significativa entre o comprometimento da atenção e a assimetria estrutural do Córtex Cingulado Anterior (CCA), mesmo sem uma associação com o volume total do CCA. Os autores sugerem que a assimetria estrutural pode levar a uma assimetria funcional. Nesse sentido, uma assimetria elevada no CCA, uma região associada à atenção, pode resultar em um desequilíbrio entre a ativação e a inibição dos hemisférios cerebrais, prejudicando o processamento de tarefas desse tipo.

3.1.2 Discussão

A análise dos trabalhos correlatos revela um consenso sobre a relevância das assimetrias inter-hemisféricas como uma característica fundamental da organização cerebral, tanto em cérebros saudáveis quanto em contextos patológicos. Os estudos de Kong et al. (2018) e Kong et al. (2022), demonstraram que tanto em indivíduos cognitivamente normais quanto em distúrbios cognitivos e neuropsiquiátricos, assimetrias na espessura cortical e área superficial são afetadas por diferentes fatores biológicos, como sexo, idade e volume intracraniano, mas não foram observados efeitos significativos de lateralidade para preferência motora. Ademais, observou-se que o cérebro humano é composto de regiões com hereditariedade significativa das características hemisféricas. Em termos de análise na DA, Long, Jiang e Zhang (2018) identificaram padrões distintos de assimetria em pacientes com CCL, um precursor da DA. Foram encontradas variações significativas nas assimetrias entre grupos CCL conversores e estáveis, especialmente em regiões associadas à linguagem e memória. Isso sugere que as assimetrias corticais, em especial o córtex entorrinal, podem ser biomarcadores significativos para diferenciar entre esses dois grupos.

Além disso, os estudos de Wachinger et al. (2015) e Wachinger et al. (2016) introduziram a técnica *BrainPrint* que analisa as assimetrias estruturais do encéfalo utilizando características das formas geométricas ao invés de dados volumétricos. Os resultados indicam que as estruturas subcorticais mostram maior assimetria em comparação com as corticais, indicando variabilidade na similaridade intra-indivíduos. Essa metodologia também revelou um aumento na assimetria cerebral associado à idade e sexo, bem como uma associação significativa entre aumento da assimetria e progressão da DA. Esses achados sugerem que a assimetria inter-hemisférica desempenha um papel importante na compreensão da DA e pode ser utilizada para monitorar a progressão da doença e possivelmente para intervenções precoces.

Os estudos de Oliveira, Poloni e Ferrari (2020) e Poloni et al. (2021) desenvolvem uma nova metodologia para a análise de assimetrias na região hipocampal em imagens de RM para classificação da DA. No primeiro estudo, Oliveira, Poloni e Ferrari (2020) utilizam

um banco de filtros log-Gabor 3D em conjunto com uma máscara binária dilatada alinhada às imagens clínicas, para identificar respostas direcionais na região hipocampal. Quatro momentos estatísticos são extraídos para cada escala da imagem filtrada, e duas abordagens são empregadas para construir vetores de características que representam a assimetria inter-hemisférica nessa região. Poloni et al. (2021) expandiram essa técnica para utilizar a correlação cruzada normalizada entre os vetores de características de cada hemisfério para avaliar a assimetria. Somado a isso, um novo índice de assimetria é proposto, e os atributos significativos identificados pelo teste ANOVA são usados para treinar e testar classificadores SVM. Os resultados demonstraram que atributos indicativos de assimetria inter-hemisférica na região hipocampal têm potencial de identificar a presença da DA, oferecendo uma abordagem complementar às técnicas tradicionais de segmentação e cálculo volumétrico.

Estudos mais recentes continuam a refinar e a contextualizar o papel das assimetrias. Korbmacher et al. (2024) introduziram o conceito de idade cerebral hemisférica, sugerindo que uma grande discrepância entre as idades estimadas para cada hemisfério pode ser um marcador de desequilíbrio na saúde cerebral. Por sua vez, o trabalho de Matsuda et al. (2025) adiciona uma importante nuance, mostrando que a associação entre assimetria e declínio cognitivo pode ser específica para determinados domínios, como a relação entre a assimetria do córtex cingulado anterior e o déficit de atenção.

Finalmente, a revisão sistemática de Ocklenburg et al. (2024) oferece uma perspectiva crítica e orientadora para futuras pesquisas. Os autores argumentam que, devido à alta variabilidade interindividual, as assimetrias podem ter valor limitado como biomarcadores diagnósticos diretos. Contudo, seu potencial é promissor para a predição de sintomas e de resposta a tratamentos, defendendo o desenvolvimento de modelos computacionais que foquem na variabilidade individual em vez de apenas na classificação de grupo.

Em resumo, as assimetrias inter-hemisféricas são relevantes tanto para a compreensão da neuroanatomia normal quanto para a análise da DA. Elas oferecem dados valiosos sobre as diferenças estruturais associadas à genética, ao envelhecimento e as condições neurodegenerativas, destacando-se como potenciais biomarcadores para o diagnóstico e monitoramento da DA. A Tabela 2 apresenta a comparação dos estudos sobre assimetria cerebral.

Tabela 2 – Comparativo dos Trabalhos Correlatos em Análise de Assimetria Cerebral.

Autor(es)	Base de Dados	Metodologia Principal	Resultados Principais
Kong et al. (2018, 2022)	ENIGMA (99 bases)	Análise de espessura e área superficial do córtex com FreeSurfer.	Mapeou assimetrias em indivíduos saudáveis e com transtornos (TEA, TDM, TOC).
Long, Jiang & Zhang (2018)	ADNI	Biomarcadores morfológicos (área, curvatura, espessura) em dados longitudinais.	Assimetria no córtex entorrinal é um biomarcador para diferenciar CCL conversor de estável.
Wachinger et al. (2015, 2016)	ADNI	BrainPrint: Análise de forma com autovalores do operador Laplace-Beltrami.	Aumento da assimetria no hipocampo e amígdala com a progressão da DA. Prediz a conversão de CCL para DA.
Oliveira, Poloni & Ferrari (2020)	ADNI, NAC	Filtros log-Gabor 3D para extrair momentos estatísticos da região hipocampal.	Atributos de assimetria hipocampal mostraram potencial para classificação da DA (Acurácia de 81,43% CN vs. DA).
Poloni et al. (2021)	ADNI, NAC	Expansão do método anterior com correlação cruzada e seleção de atributos com ANOVA.	Acurácia de 82,59% (CN vs. DA) usando apenas atributos de assimetria do hipocampo.
Korbmacher et al. (2024)	UK Biobank	Idade cerebral hemisférica (HBA) com RM multimodal.	Discrepância entre HBAs pode ser um biomarcador de desequilíbrio na saúde cerebral.
Ocklenburg et al. (2024)	Revisão Sistemática	Revisão sobre o valor clínico das assimetrias cerebrais.	Assimetria tem valor limitado para diagnóstico, mas é promissora para prever sintomas e resposta a tratamentos.
Matsuda et al. (2025)	NCGG-SGS	Análise volumétrica na população idosa.	Assimetria estrutural está associada à alterações em testes neuropsicológicos.

Fonte: elaborado pelo autor.

3.2 Pontos salientes 3D

No contexto da Visão Computacional, um ponto saliente é uma região local em uma imagem que se destaca devido à sua singularidade em comparação à sua vizinhança imediata (LOWE, 2004), sendo estável mesmo sob transformações como mudanças no ponto de vista, iluminação, escala, desfoque e compressão (LIU; XU; WANG, 2021). Os pontos salientes são cruciais para aplicações como reconhecimento de objetos ou cenas, correspondência e rastreamento de movimento, e são essenciais na etapa de detecção e descrição em processos de corregristo de imagens (LIU; XU; WANG, 2021). No contexto desta pesquisa, pontos salientes no cérebro não precisam, necessariamente, representar uma estrutura anatômica específica. Assim, variações locais de intensidade, textura ou outras métricas, podem gerar pontos salientes.

A próxima seção visa avaliar os técnicas e descritores de pontos salientes comumente utilizados para imagens 3D, com foco em técnicas aplicadas às imagens médicas.

3.2.1 Trabalhos correlatos

A técnica *Scale Invariant Feature Transform* (SIFT) desenvolvida por Lowe (2004) calcula uma aproximação do filtro Laplaciano em imagens 2D em diferentes escalas utilizando diferenças de Gaussianas. Essa técnica permite a detecção de pontos salientes

em diferentes escalas e parcialmente invariantes a variações de intensidade. Ademais, os pontos obtidos são invariantes às rotações uma vez que cada ponto possui uma orientação dominante calculada com o histograma de gradientes orientados do filtro Laplaciano. Com isso, também obtêm-se um descritor localizado em cada ponto que pode ser utilizado para identificar os objetos presentes na cena. Nesse contexto, o trabalho Rister, Horowitz e Rubin (2017) foi o primeiro estudo que estendeu a técnica SIFT para $\mathbb{R}^n$ considerando as orientações do histograma de gradientes orientados sem serem afetados por anisotropia. Para atingirem esses objetivos, os autores propuseram modificações nas etapas de detecção dos pontos salientes e nos descritores. Diferentemente de trabalhos anteriores, que utilizavam uma vizinhança ℓ^∞ $(3^n - 1)$ para a detecção dos pontos de máximos e mínimos locais, os autores, inspirados por trabalhos de Moore e von Neuman em autômatos celulares, utilizaram uma vizinhança ℓ^1 conectada em $2n$. Para definir uma orientação local à cada ponto, que seja isotrópica e aplicável em $\mathbb{R}^n$, os autores utilizaram um tensor estruturante (*structure tensor* - ST), isto é, a correlação entre os gradientes do Laplaciano, aproximado pela diferença de Gaussianas dada por

$$K = \int \omega(\mathbf{x}) \nabla I(\mathbf{x}) (\nabla I(\mathbf{x}))^T d\mathbf{x}, \quad (54)$$

em que $\nabla I(\mathbf{x})$ é o gradiente da imagem I na posição $\mathbf{x}$, e $\omega(\mathbf{x})$ é uma Gaussiana centrada no ponto saliente. Os pontos salientes considerados instáveis são determinados calculando o ângulo entre o gradiente da imagem $\mathbf{d}$ e os autovetores $\mathbf{q}_i$ do ST, sendo

$$\cos(\theta_i) = \frac{\mathbf{q}_i^T \mathbf{d}}{\|\mathbf{q}_i\| \|\mathbf{d}\|}, \quad (55)$$

a fórmula que determina o ângulo entre esses vetores, e

$$\max_i \left(\left| \frac{\lambda_i}{\lambda_{i+1}} \right| \right) > \beta, \quad (56)$$

a equação que determina a relação entre os autovalores λ_i ordenados em ordem crescente e β é o limiar utilizado para determinar os pontos instáveis que devem ser descartados, definido como 0,9 pelos autores.

Para testar esse detector, Rister, Horowitz e Rubin (2017) utilizaram imagens longitudinais de um paciente com esclerose múltipla, sendo a imagem obtida na primeira sessão do cérebro sem lesões. Pontos salientes foram detectados e seus descritores extraídos de ambas imagens, os pontos salientes da imagem sem lesões foram definidos como pontos verdadeiros e comparados, utilizando a técnica de casamento proposta por Lowe (2004), para determinar os pontos detectados na posição correta. Um ponto detectado na imagem com lesões foi considerado verdadeiro positivo quando possuiu um casamento na imagem sem lesões e a distância entre os pontos foi inferior à um limiar ε . A vizinhança ℓ^∞ proposta pelos autores obteve 166 casamentos totais, sendo 154 casamentos verdadeiros com distância $\varepsilon = 5$ mm e 129 com $\varepsilon = 2$ mm. Enquanto que a vizinhança ℓ^1 obteve 2818

casamento totais, sendo 2695 casamentos verdadeiros com distância $\varepsilon = 5$ mm e 2229 com $\varepsilon = 2$ mm.

Korb e Ferrari (2020) utilizaram a técnica SIFT de Rister, Horowitz e Rubin (2017) para detectar pontos em imagens de RM com o objetivo de auxiliar o correto registro e posicionamento de modelos geométricos (malhas) em tarefas de segmentação automática, com enfoque para a segmentação dos hipocampus. Antes da aplicação em imagens de RM, Korb e Ferrari (2020) avaliaram a técnica SIFT em imagens sintéticas 3D contendo estruturas de alta curvatura (*tip-like*) com diferentes variações de translação, rotação, intensidade, flexão, e com a adição de ruído Gaussiano e *bias field*. Essas imagens foram geradas usando o modelo paramétrico proposto por Wörz e Rohr (2006), com resolução isotrópica de 1mm, tamanho de $33 \times 33 \times 33$ voxels, e localização da posição de máxima curvatura determinada matematicamente. O detector SIFT falhou em detectar pontos salientes nos vértices de um cubo e nas imagens *tip-like* que possuíam isotropia em duas das três dimensões. Dessa forma, Korb e Ferrari (2020) modificaram o detector eliminando o uso da Equação 55 para reduzir a quantidade de *outliers* e substituíram a Equação 56, por

$$T_1 < \left| \frac{\lambda_i}{\lambda_{i+1}} \right| < T_2, \quad (57)$$

em que $T_1 = 0,1$ é utilizado para garantir a existência de variância em todas as três direções, e $T_2 = 1,0$ serve como um filtro anisotrópico que descarta cantos com curvatura isotrópica.

Utilizando o SIFT (RISTER; HOROWITZ; RUBIN, 2017) e sua versão modificada, Korb e Ferrari (2020) aplicaram o detector para a obtenção de pontos salientes em imagens clínicas de RM e na imagem de referência do atlas NAC. Seja X e Y duas imagens, a correspondência entre um descritor $\mathbf{x} \in X$ e $\mathbf{y} \in Y$ é obtida utilizando a seguinte equação:

$$g(\mathbf{x}, S_2) = \frac{d(\mathbf{x}, \mathbf{y}_1)}{d(\mathbf{x}, \mathbf{y}_2)}, \quad (58)$$

em que $\mathbf{x}$ é um descritor pertencente à imagem X , $\mathbf{y}_1$ é o descritor pertencente à imagem Y mais próximo de $\mathbf{x}$, $\mathbf{y}_2$ é o descritor pertencente à Y e segundo mais próximo de $\mathbf{x}$, e d é a distância Euclidiana entre eles (LOWE, 2004; RISTER; HOROWITZ; RUBIN, 2017). Após o casamento dos pontos salientes detectados na imagem clínica e na imagem de referência, esses pontos foram utilizados para estimar uma transformação *B-spline* entre as imagens. A transformação resultante, representada como um campo de deformação (*Displacement Field* - DF), é aplicada aos vértices do modelo geométrico na imagem de referência para obter o posicionamento dos novos vértices posicionados sob a imagem clínica, que define a segmentação. Korb e Ferrari (2020) utilizaram o teste de Mann-Whitney (U) (CORDER; FOREMAN, 2014) para verificar se a diferença na média do coeficiente Dice é estatisticamente significativo. Os resultados obtidos pelos autores são apresentados na Tabela 3.

Tabela 3 – Valores médios do coeficiente de Dice em 129 imagens da base ADNI.

	Hipocampo esquerdo			Hipocampo direito		
	Média	U	p	Média	U	p
Afm	$0,49 \pm 0,12$	-	-	$0,48 \pm 0,12$		
SIFT 3D	$0,42 \pm 0,17$	11286	0,0007	$0,46 \pm 0,17$	9418	0,6345
SIFT 3D modificado	$0,55 \pm 0,09$	6526	0,0001	$0,53 \pm 0,12$	7118	0,0019

Fonte: Korb e Ferrari (2020).

No trabalho de Souza, Poloni e Ferrari (2020), os autores expandiram a técnica de Fauqueur, Kingsbury e Anderson (2006) para aplicação em imagens volumétricas 3D. Utilizando a técnica DT-CWT para imagens volumétricas, conforme descrita na Seção 2.6.4, são geradas 28 sub-bandas direcionais por escala, que são consolidadas em um único mapa de energia 3D. Este detector de pontos salientes 3D é utilizado para estimar uma transformação deformável, aplicada aos vértices de malhas triangulares de estruturas cerebrais, facilitando o posicionamento inicial dessas estruturas. A técnica foi testada em imagens volumétricas sintéticas representando estruturas de alta curvatura, semelhantes às encontradas em imagens de RM cerebral. Os dados sintéticos foram gerados por um modelo paramétrico definido por Wörz e Rohr (2006), totalizando 648 imagens distribuídas em 12 grupos definidos por diferentes níveis de degradação causada por ruído gaussiano e *bias field* (BF). Primeiramente, foi realizada a otimização dos parâmetros com busca em grade, resultando nos seguintes valores: $\alpha = 1$, $\beta = 0.05$, limiar $T_E = 0.5$, tamanho da vizinhança de máximo igual à $3 \times 3 \times 3$, e $s = 3$ escalas. Em seguida, para avaliar a reprodutibilidade do detector na presença de ruídos, os autores detectaram os pontos salientes na imagem de referência da NAC e em nos subgrupos das imagens afetadas por deformações da imagem de referência. Os mapas de deformações foram usados para determinar qual deveria ser a posição correta dos pontos detectados nas imagens deformadas e, por conseguinte, contar os pontos não detectados (falso negativo - FN) e os detectados à mais de 10 voxels da localização correta (falso positivo - FP), enquanto que a imagem de referência da NAC, não afetada por ruídos, serve como referência dos pontos verdadeiros (verdadeiro positivo - VP). Os valores de FP e FN obtidos são apresentados na 4. Portanto, com exceção da influência negativa do BF, o detector é capaz de detectar pontos salientes corretamente.

De maneira análoga, visando o uso de pontos de referência em estruturas anatômicas para auxiliar no correto registro com transformações deformáveis de imagens, Grewal et al. (2023) apresentam uma rede neural convolucional profunda (*Deep Convolutional Neural Network* - DCNN) para a detecção e correspondência automática de pontos salientes em imagens 3D. As técnicas convencionais de correspondência desses pontos envolvem três etapas: (a) a detecção independente de pontos salientes em cada imagem; (b) a extração de um vetor de atributos, ou seja, um descritor, para cada ponto saliente detectado; (c)

Tabela 4 – Valores de FP e FN sob a influência de ruídos.

	$BF = 0\%$		$BF = 20\%$		$BF = 40\%$	
	FP	FN	FP	FN	FP	FN
$\sigma_r = 0\%$	3	0	3	29	1	45
$\sigma_r = 1\%$	4	0	4	29	6	45
$\sigma_r = 2\%$	4	0	4	27	15	45
$\sigma_r = 3\%$	4	0	4	29	67	46

Fonte: Souza, Poloni e Ferrari (2020)

a correspondência (casamento) dos descritores usando um algoritmo com uma métrica para calcular a similaridade entre eles. A metodologia proposta por Grewal et al. (2023) substitui essas etapas por uma rede neural modular, em que a saída de uma rede é a entrada da próxima, configurando uma proposta *end-to-end*. A arquitetura consiste em uma rede siamesa com três módulos: (a) duas DCNN com *kernels* 3D e pesos compartilhados, (b) uma camada de amostragem e (c) um módulo de casamento de descritores. Essa arquitetura é uma extensão do detector 2D de Grewal et al. (2020), que por sua vez é derivada da rede UNet (RONNEBERGER; FISCHER; BROX, 2015) desenvolvida para tarefas de segmentação automática em imagens biomédicas. O objetivo da DCNN é calcular o mapa de probabilidade $\hat{p}_i^x(x \in \{target, source\})$ de cada voxel i ser um ponto saliente, sendo I_{source} a imagem a ser avaliada e I_{target} uma imagem de referência. Adicionalmente, a camada de amostragem tem como função primordial criar um par de vetores de atributos $f_i^{I_x}$ para cada i -ésimo ponto saliente detectado. Por fim, a última camada realiza o casamento dos descritores $(f_i^{I_{target}}, f_j^{I_{source}})$ para todos os $i = 1, 2, \dots, K$ em I_{target} e $j = 1, 2, \dots, K$ em I_{source} , determinando a probabilidade de dois pares de descritores representarem a mesma estrutura anatômica localizada no ponto saliente.

A DCNN de Grewal et al. (2023) foi treinada usando 1.335 imagens de tomografia computadorizada (TC), e foi testada em 336 imagens, sendo que 10 imagens de CT foram utilizadas para validação dos hiperparâmetros. Para avaliar a detecção e casamentos dos pontos salientes, foram utilizados dois conjuntos de testes (deformações simuladas e clínicas) com dados distintos de pacientes que receberam tratamento de radiação para câncer cervical. O conjunto de deformações simuladas incluiu 121 imagens de TC, sendo I_{target} as imagens originais e I_{source} as imagens que sofreram transformações deformáveis, formando os pares de entrada da rede DCNN. O conjunto de deformações clínicas foi composto por 11 pares imagens de TC capturadas no início do tratamento (I_{target}) e em exames de retorno (I_{source}), com 29 pontos salientes identificados manualmente por um especialista em cada par de imagens, representando estruturas anatômicas. Dessa forma, em ambos os conjuntos, a posição exata de um ponto saliente na imagem I_{source} pode ser determinada a partir de sua localização na imagem I_{target} . A diferença entre a posição esperada e a calculada pela DCNN foi usada como métrica para avaliar o detector. Para ambos os conjuntos de deformações, após ajustes dos hiperparâmetros, o melhor resultado

alcançado foi a detecção de 90% dos pontos salientes com um erro inferior a 1 voxel (2 mm) de distância da posição esperada.

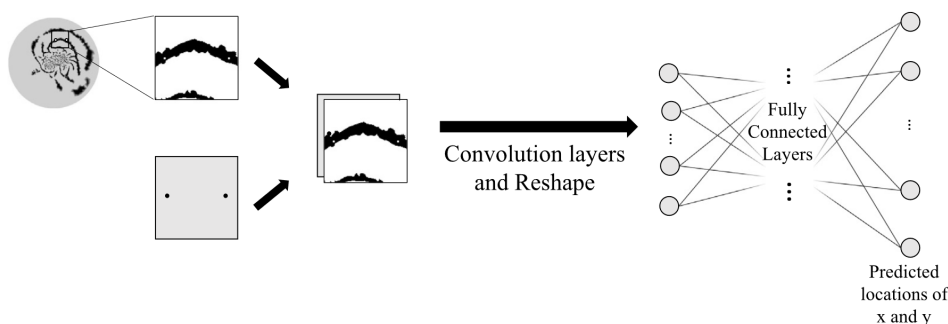
Chauvin et al. (2022) apresentaram uma nova métrica para avaliar a distância entre conjuntos de descritores extraídos de imagens distintas. O objetivo da pesquisa foi avaliar a relação entre o fenótipo cerebral humano e o genótipo subjacente, questionando até que ponto o cérebro é influenciado pela genética. Os autores investigaram como a variação da imagem de RM do cérebro se encontra localmente associada ao genótipo, como entre parentes próximos ou grupos definidos por fatores genéticos sutis, como ancestralidade racial ou sexo. Para isso, o estudo introduz uma nova medida de distância para conjuntos de pontos salientes 3D, generalizando o índice de Jaccard para considerar a equivalência parcial entre conjuntos suaves. Os autores utilizaram a técnica 3D SIFT-Rank (TOEWS; WELLS, 2013) para detectar pontos salientes no cérebro em imagens volumétricas de RM. A distância entre os pontos salientes é determinada utilizando um *kernel* para medir a similaridade entre seus respectivos descritores, e um novo *kernel* é introduzido para normalizar a diferença entre pontos detectados em diferentes escalas da pirâmide de imagens. O método proposto é eficiente em indexação dos pontos salientes e escalável para grandes conjuntos de dados.

Esse novo índice foi avaliado em 1010 imagens de 434 famílias, da base de dados *Human Connectome Project* (HCP)⁵. Após o pré-processamento para alinhamento espacial das imagens e remoção da caixa craniana, o detector extraiu aproximadamente 1400 pontos salientes em cada imagem. Utilizando processamento em GPU, os *k*-vizinhos mais próximos são calculados entre todos os descritores da base de dados. O teste de Kolmogorov-Smirnoff (KS)(BERGER; ZHOU, 2014) indicou diferenças significativas entre as distribuições, exceto entre irmãos dizigóticos e irmãos de pais diferentes. A distância de Jaccard modificada foi capaz de identificar pares atípicos em suas distribuições, destacando-se pela precisão da distância entre os vizinhos mais próximos na previsão de famílias autodeclaradas. Sujeitos não relacionados compartilham quantidades nominais de informações genéticas, com semelhanças sutis devido a fatores demográficos. Esperava-se que a distância cerebral total fosse menor para pares da mesma raça e sexo e maior para diferentes raças e sexos. Entretanto, as distribuições de distâncias permaneceram virtualmente idênticas em toda a faixa etária dos sujeitos da base HCP, indicando que a diferença de idade não é um fator de confusão significativo neste grupo relativamente jovem e saudável. Em suma, a metodologia proposta utilizando o índice de Jaccard modificado demonstrou ser eficaz em identificar a similaridade entre os descritores extraídos de pontos salientes detectados em indivíduos pertencentes à mesma família.

O artigo de Guo et al. (2023) apresenta uma nova técnica para detecção de pontos salientes e correção de superfícies corticais cerebrais que emprega uma combinação de geometria quase-conforme e redes neurais convolucionais (*Convolutional Neural Network*

⁵ Disponível em: <https://www.humanconnectome.org/> Acesso em: 01 de dez, 2023.

Figura 8 – Diagrama da etapa de normalização, redimensionamento e detecção dos pontos salientes localizados no sulco limitado por dois pontos.



Fonte: Guo et al. (2023).

- CNN). Esse método tem como objetivo detectar pontos salientes localizados em sulcos cerebrais, delimitados por dois pontos específicos. A abordagem adotada envolve a transformação das coordenadas de malhas triangulares para um disco unitário 2D via mapeamento quase-conforme. A metodologia utiliza três redes neurais, representando três etapas distintas: (i) a rede para a detecção de pontos salientes localizados em sulcos do encéfalo (*landmark detection network* - LD-Net), (ii) a rede para predição dos coeficientes de Beltrami (*coefficient prediction network* - CP-Net) e (iii) a rede para solução da equação de Beltrami em disco (*disk Beltrami solver network* - DBS-Net). Diferentemente de técnicas anteriores, o detector de pontos salientes proposto por Guo et al. (2023) utiliza, além de uma malha triangular, dois pontos - um inicial e um final - que delimitam um sulco cerebral. Para cada malha tridimensional, calcula-se a curvatura média em cada vértice, seguida pela parametrização da superfície para um disco unitário 2D, conforme a técnica de Choi e Lui (2015). Em seguida, a malha transformada é rotacionada e reescalada, garantindo que a posição xy dos pontos extremos estejam sempre na mesma orientação horizontal e contidos em uma janela de tamanho constante. Nesta etapa, os pixels da janela são normalizados, formando uma imagem binária de dois canais, que inclui as localizações dos pontos extremos. Regiões claras e escuras visualizadas após a binarização correspondem, respectivamente, aos giros (cristas) e sulcos da superfície cortical do encéfalo. Essa imagem é então inserida na LD-Net, que é composta por camadas convolucionais, uma etapa de redimensionamento dos atributos e camadas densamente conectadas, resultando na saída das coordenadas xy dos pontos salientes detectados. A Figura 8 ilustra as etapas de processamento do disco unitário e arquitetura da LD-Net.

Para avaliar essa técnica, foram extraídas 24 superfícies com o software *FreeSurfer* em imagens de RM da base de dados *Open Access Series of Imaging Studies* (Open Access Series of Imaging Studies (OASIS))⁶. Foram selecionados seis sulcos, e as posições dos pontos de início e término foram calculadas para serem utilizadas como entrada na rede. Os autores manualmente rotularam diversos pontos salientes nos sulcos cerebrais

⁶ Disponível em: <http://www.oasis-brains.org/> Acesso em: 05 dez, 2023

em algumas imagens de referência. Para essas imagens, os dois pontos extremos foram definidos aleatoriamente, parametrizados, normalizados e utilizados para treinar a LD-Net. A rede foi então avaliada pela comparação dos pontos detectados automaticamente com aqueles selecionados por um especialista. Os resultados obtidos apresentaram uma distância média e desvio padrão de 0,0089 e 0,0052, respectivamente, indicando a precisão do método proposto.

Em um estudo focado no diagnóstico de doenças cerebrais por meio de imagens de RM, Li, Ji e Sun (2024) propuseram uma abordagem baseada em aprendizado profundo para a classificação do TEA e da DA. Para a análise, os autores utilizaram duas bases de dados de acesso público. A primeira, proveniente da Autism Brain Imaging Data Exchange (ABIDE), contém imagens de RM ponderadas em T1 de 213 indivíduos com TEA e 342 controles com desenvolvimento típico. A segunda base, da ADNI, era composta por imagens de 308 pacientes com DA e 399 controles saudáveis.

A metodologia desenvolvida, denominada *Landmark-based Multi-Instance Conv-Transformer* (LD-MILCT), combina CNNs e a arquitetura *Vision Transformer* (ViT) em um *framework* de aprendizado por múltiplas instâncias. O método primeiramente identifica, de maneira analítica, pontos anatômicos que apresentam diferenças estatisticamente significativas entre os grupos. Em seguida, sub-volumes 3D são extraídos ao redor desses pontos e processados por blocos de CNNs para capturar características estruturais locais. Subsequentemente, um módulo Transformer é empregado para modelar as correlações entre as informações locais e globais de todas as regiões de interesse, superando uma limitação comum de aprendizado por múltiplas instâncias, que ignoram a interdependência entre diferentes áreas cerebrais. Os resultados do estudo demonstraram a superioridade do modelo proposto em comparação com outros métodos de aprendizado profundo e baseados em pontos. Na classificação da DA, o modelo LD-MILCT alcançou uma acurácia (ACC) de 91,0% e um F1-score de 0,903. Para a classificação do Transtorno do Espectro Autista, uma tarefa considerada mais desafiadora devido à sutileza das anomalias estruturais, a mesma arquitetura obteve uma acurácia de 70,1% e um F1-score de 0,333. Esses resultados, na tarefa de classificação da DA, ressaltam o potencial da combinação de extração de características locais baseada em marcos com a capacidade dos Transformers de modelar relações de longa distância para o diagnóstico auxiliado por computador de doenças neurológicas.

3.2.2 Discussão

Os estudos revisados demonstram avanços significativos no desenvolvimento e aplicação de detectores de pontos salientes 3D, especialmente em imagens de ressonância magnética, como pode ser observado na Tabela 5. A técnica SIFT Lowe (2004) ampliada para 3D por Rister, Horowitz e Rubin (2017) destaca-se pela sua capacidade de detectar pontos salientes em diferentes escalas e sua invariância a rotações. Esta técnica foi

adaptada com sucesso para imagens de RM, mostrando a eficácia na detecção de pontos correspondentes em imagens com e sem lesões, o que sugere seu potencial na identificação de mudanças estruturais relacionadas a doenças. No estudo de Korb e Ferrari (2020), a técnica SIFT foi empregada para detectar pontos salientes em imagens de RM com o intuito de facilitar o corregristo e o posicionamento de modelos geométricos para realizar a segmentação automática de estruturas anatômicas cerebrais, focando na segmentação dos hipocampos. O estudo avaliou a eficácia da técnica SIFT de Rister, Horowitz e Rubin (2017) e constatou que essa técnica apresentou limitações ao falhar na detecção de pontos em certas estruturas, que foram corrigidas com modificações nos cálculos dos pontos salientes estáveis. Esta versão modificada foi aplicada em imagens clínicas de RM e em uma imagem de referência, utilizando uma equação para estabelecer correspondências entre descritores das imagens e, em seguida, uma transformação *B-spline* foi estimada para posicionar os novos vértices do modelo geométrico na imagem clínica. Os resultados obtidos demonstraram uma melhora no posicionamento dos modelos geométricos quando utilizados com o detector modificado.

A metodologia baseada na DT-CWT, proposta por Fauqueur, Kingsbury e Anderson (2006) e expandida por Souza, Poloni e Ferrari (2020) para imagens 3D, também demonstra robustez em detectar pontos salientes em imagens volumétricas, incluindo imagens de RM cerebral. Este método revelou-se eficaz mesmo na presença de ruídos, o que é crucial para análises clínicas.

No estudo de Chauvin et al. (2022), os autores investigaram a relação entre o genótipo e o fenótipo cerebral usando uma nova medida de distância para conjuntos de pontos salientes, com foco na previsão de relações familiares a partir de dados de imagens médicas volumétricas. Utilizando uma abordagem baseada em *kernels* que combina aparência e geometria local ao redor de cada ponto saliente, o estudo revelou uma distinção clara na similaridade do índice entre as imagens de indivíduos da mesma família em relação às imagens de indivíduos não relacionados. Além disso, a distância de Jaccard modificada mostrou-se eficaz na previsão de relações familiares e também foi adaptada para prever rótulos de grupo, como sexo, usando conjuntos formados pela união dos membros do grupo, alcançando alta precisão. Essa abordagem demonstra que a similaridade entre conjuntos de pontos salientes e descritores pode ser utilizada com eficácia para avaliar imagens de RM do cérebro.

Os trabalhos de Grewal et al. (2023) e Guo et al. (2023) introduzem abordagens utilizando redes neurais convolucionais para a detecção de pontos salientes em imagens de RM e TC, respectivamente. Essas técnicas, ao contrário das anteriores, empregam redes para processar as imagens ou malhas triangulares e identificar pontos salientes. Por fim, o trabalho de Li, Ji e Sun (2024) utiliza pontos salientes anatômicos, identificados por uma abordagem que busca diferenças estatisticamente significativas entre grupos, para guiar a extração de *patches* que servirão de entrada para uma rede neural.

Em resumo, detectores de pontos salientes 3D oferecem um meio robusto e preciso para identificar as localizações de características estruturais específicas do cérebro. Eles têm o potencial de auxiliar significativamente na identificação de biomarcadores de assimetrias inter-hemisféricas, e auxiliar o diagnóstico de diversas condições neurológicas e psiquiátricas.

Tabela 5 – Comparativo dos Trabalhos Correlatos em Detecção de Pontos Salientes 3D.

Autor(es)	Base de Dados	Metodologia Principal	Resultados Principais
Rister et al. (2017)	RM longitudinal de esclerose múltipla.	Extensão do SIFT para $\mathbb{R}^n$, propondo uma vizinhança ℓ^1 e um tensor estruturante para orientação isotrópica.	A vizinhança ℓ^1 obteve 2229 casamentos verdadeiros ($\varepsilon = 2$ mm), superando significativamente a vizinhança ℓ^∞ (129).
Korb et al. (2020)	Imagens sintéticas e 129 RMs da ADNI.	Modificação do SIFT 3D para melhorar a detecção em estruturas anisotrópicas, visando a segmentação do hipocampo.	O método modificado alcançou um Dice de 0,55 no hipocampo esquerdo, superior à versão original do SIFT 3D (0,42).
da Silveira Souza et al. (2020)	Imagens sintéticas e atlas NAC.	Extensão de um detector baseado em DT-CWT para 3D, gerando um mapa de energia a partir de 28 sub-bandas.	O detector mostrou-se robusto a ruído Gaussiano, mas teve o desempenho degradado pela presença de <i>bias field</i> .
Grewal et al. (2023)	Imagens de Tomografia Computadorizada (TC).	DCNN <i>end-to-end</i> baseada na U-Net para detecção e correspondência automática de pontos salientes.	Detectou 90% dos pontos com erro inferior a 1 voxel (2 mm) em testes com deformações simuladas e clínicas.
Chauvin et al. (2022)	HCP	Nova métrica de distância (Jaccard generalizado) para conjuntos de descritores SIFT-Rank 3D.	A métrica foi eficaz em quantificar a similaridade fenotípica cerebral, prevenindo relações familiares com precisão.
Guo et al. (2023)	OASIS	CNN (LD-Net) para detectar pontos em sulcos cerebrais a partir de malhas corticais mapeadas para um disco 2D.	Alta precisão na detecção, com distância média de 0,0089 em relação aos pontos rotulados manualmente.
Li et al. (2024)	ABIDE, ADNI	LD-MILCT: Framework que combina CNN e Transformer para classificar sub-volumes extraídos de pontos anatômicos.	Acurácia de 91,0% na classificação de CN vs. DA e 70,1% na classificação de TEA vs. CN.

Fonte: elaborado pelo autor.

3.3 Progressão do CCL para a DA

O CCL representa um estágio de transição entre o envelhecimento cognitivo normal e o quadro demencial que define a DA (Alzheimer’s Association, 2023). Indivíduos diagnosticados com CCL apresentam risco elevado de conversão para a DA, contudo, a taxa de progressão pode variar, tornando a identificação de biomarcadores prognósticos uma prioridade clínica e de pesquisa. Nesse contexto, a RM estrutural surgiu como uma ferramenta não invasiva essencial, permitindo a detecção e quantificação de alterações morfológicas cerebrais, como a atrofia em regiões-chave associadas à patologia da DA (SCHELTENS et al., 1992).

Para analisar os dados tri-dimensionais de neuroimagem e extrair padrões que sejam preditivos da conversão, estudos recentes tem se voltado para o uso de modelos de aprendizado profundo, especialmente CNNs, que são capazes de aprender representações de características diretamente a partir das imagens do encéfalo. Esta seção revisa trabalhos recentes que aplicam essas técnicas para a predição da progressão do CCL para a DA, detalhando suas metodologias, resultados e lacunas.

3.3.1 Trabalhos correlatos

Uma revisão sistemática realizada por Niyas e Thiagarajan (2023) analisou 42 estudos focados na predição da conversão de CCL para a DA utilizando algoritmos de aprendizado de máquina em dados longitudinais. O estudo investigou criticamente diversos aspectos das pesquisas publicadas, incluindo as modalidades de dados empregadas, as técnicas de extração e seleção de atributos, e os modelos de aprendizado de máquina e profundo utilizados. O levantamento revelou um uso predominante de dados de neuroimagem, como a RM e a PET, com diversos estudos aplicando técnicas de extração de atributos via redes neurais profundas, especialmente CNNs, para processar imagens do cérebro inteiro e ROIs específicas.

A análise comparativa dos resultados indicou que abordagens multimodais geralmente apresentam um desempenho superior em comparação com modelos que utilizam apenas uma modalidade. Assim, o melhor desempenho identificado na revisão foi obtido pela combinação de RM funcional e RM estrutural, que atingiu uma acurácia de 96%. Apesar dos avanços, Niyas e Thiagarajan (2023) destacam desafios de pesquisa significativos que ainda persistem no campo. Entre eles, a identificação de biomarcadores de neuroimagem precisos para a conversão de CCL para DA, a dificuldade em generalizar os modelos para diferentes populações e contextos hospitalares, e a complexidade de prever a conversão em curtos intervalos de tempo (e.g., 6 a 12 meses).

Em um estudo de Luo et al. (2023), os autores propõem o método de aprendizado profundo *Class Activation Attention Transfer* (CAAT) para prever a conversão de CCL para a DA a partir de uma única imagem de RM da visita inicial. Esse método permite a visualização de *Class Activation Maps* (CAMs) que realçam regiões canônicas da patologia de DA. Os autores utilizaram a base ADNI, com 593 indivíduos diagnosticados com CCL. O grupo CCLp foi composto por indivíduos que converteram para a DA em até 36 meses, enquanto o grupo CCLe foi composto por aqueles que mantiveram o diagnóstico de CCL no mesmo período. Nos experimentos, o modelo proposto pelos autores obteve acurácia de $74,61\% \pm 0,002$, F1-score de $0,75 \pm 0,002$ e AUC de $0,78 \pm 0,003$, enquanto que o modelo de controle, uma 3D-CNN, obteve acurácia de $70,84\% \pm 0,001$, F1-score de $0,71 \pm 0,003$ e AUC de $0,758 \pm 0,003$. Dessa forma, os resultados obtidos indicam ganhos positivos do modelo proposto, que dispensa a utilização da segmentação de ROIs e permite uma maior interpretabilidade via CAMs.

Em um estudo subsequente com o mesmo conjunto de dados, Luo et al. (2024) aprimoraram o modelo anterior ao implementar um mecanismo de dupla atenção para fundir as características da imagem com metadados neurocognitivos (testes de avaliação neuropsicológica) que acompanham as imagens da base ADNI. Essa modificação teve como objetivo guiar o foco da rede para as regiões cerebrais mais discriminativas com base nas informações dos testes. Os resultados experimentais demonstraram uma melhora significativa de desempenho com a nova abordagem, alcançando uma acurácia de $81,34\% \pm 0,001$,

F1-score de $0,82 \pm 0,001$ e AUC de $0,874 \pm 0,001$. Um estudo de ablação confirmou a contribuição positiva de ambos os mecanismos de atenção para o desempenho final do modelo, evidenciando que a fusão de características de imagem com metadados neurocognitivos é uma estratégia eficaz para a predição da progressão do CCL.

Wang et al. (2024) realizaram uma revisão sistemática de 16 estudos, envolvendo um total de 9.290 participantes, como objetivo de fornecer um panorama sobre os modelos preditivos do risco de demência por DA em idosos com CCL. O estudo revelou que a maioria dos modelos (11 de 16) utilizava abordagens de aprendizado de máquina, enquanto quatro empregavam métodos de modelagem tradicional. A análise da frequência dos preditores identificou que idade, o Questionário de Atividades Funcionais (*Functional Activities Questionnaire* - FAQ) e o Mini Exame do Estado Mental (*Mini-mental State Examination* - MMSE) foram os mais comuns, juntamente com biomarcadores de imagem como volume hipocampal e espessura cortical.

A meta-análise de 15 dos modelos preditivos resultou em uma AUC combinada de 0,86, indicando, em geral, um bom poder discriminativo. Notavelmente, a análise de subgrupos demonstrou que os modelos baseados em aprendizado de máquina alcançaram um desempenho superior, com uma AUC de 0,89, em comparação aos modelos de regressão tradicionais, que obtiveram uma AUC de 0,77. Esses resultados sugerem um potencial maior das técnicas de aprendizado de máquina para capturar as relações nos dados que melhor preveem a progressão da doença.

Apesar do desempenho promissor, o estudo de Wang et al. (2024) aponta para deficiências metodológicas significativas na literatura existente, resultando em um risco de viés alto ou incerto para quase todos os modelos avaliados. As principais críticas incluem a falta de validação externa, visto que apenas dois dos 16 modelos foram validados em bases de dados independentes, além de problemas recorrentes como tamanho amostral insuficiente, tratamento inadequado de dados faltantes e transparência na apresentação dos modelos. Nesse sentido, os autores concluem que nenhum modelo atual pode ser recomendado para uso clínico e reforçam a necessidade de futuras pesquisas se concentrarem no aprimoramento do rigor metodológico, na validação externa e na apresentação transparente para desenvolver ferramentas preditivas que sejam robustas e generalizáveis.

Rahim et al. (2025) propuseram a utilização de múltiplas visualizações 2D de imagens de RM para prever a conversão de CCL para DA. Os autores desenvolveram um modelo de *ensemble* profundo que integra dados dos planos axial, coronal e sagital de volumes de RM 3D, com o objetivo de capturar as características morfológicas sutis da progressão da DA em regiões cerebrais vulneráveis. A metodologia proposta envolve a extração de características de cada plano de RM via CNNs como EfficientNet-B7, ConvNext e DenseNet-121. Somado a isso, um módulo de atenção foi aplicado às características extraídas para minimizar possíveis informações redundantes entre os modelos. Em seguida, esse conjunto de características otimizado foi processado por três classificadores: Multilayer Percep-

tron (MLP), Long-Short Memory Transformer (LSMT) e Multi-Head Attention (MHA), cujos hiperparâmetros foram ajustados via otimização bayesiana. A fim de aumentar a interpretabilidade do modelo, a técnica Grad-CAM foi utilizada para gerar mapas de atenção que identificam visualmente as regiões cerebrais críticas que influenciaram as decisões do modelo.

A metodologia proposta foi avaliada em 192 indivíduos da base ADNI, sendo 96 pertencentes ao grupo CCL estável (CCLe), 81 do grupo CCL progressivo (CCLp) e 15 do grupo DA. Esses indivíduos foram divididos em dois grupos para a classificação: (i) CCLe e (ii) CCLp + DA. Como critério de rotulação, os autores consideraram CCLe aqueles diagnosticados como CCL na visita inicial e que permaneceram nesse grupo nos 48 meses seguintes. Já o grupo CCLp foi composto por indivíduos diagnosticados como CCL na visita inicial e que progrediram para o diagnóstico de DA entre 18 e 48 meses. Com isso, o modelo de *ensemble* proposto, combinado com o mecanismo de atenção e o classificador MHA, alcançou $91,77\% \pm 3,14$ de acurácia e AUC de $92,87\% \pm 3,58$. Os mapas de atenção revelaram padrões de danos teciduais consistentes em regiões cerebrais notórias por serem afetadas no início da DA, como o hipocampo e o córtex entorrinal.

O trabalho de Liu et al. (2025) propõe o AFSNet, uma CNN 3D guiada por atenção para previsão precoce da progressão de CCL para DA em imagens de RM longitudinais. Dentre as principais contribuições, é possível destacar os três mecanismos: (i) *longitudinal lesion feature selection* (LLFS), que calcula diferenças entre mapas de atributos T1 e T2, e aprende, para cada canal, pesos para realçar variações associadas à doença; (ii) *time-flow integration*, que une o mapa de atributos T2 com o mapa de diferenças (T1 – T2) para codificar tendências globais e locais; e (iii) *disease trend attention mechanism*, um módulo tipo SEnet que prioriza canais relevantes à progressão. Os autores também empregam visualização (Layer-CAM em 3D) para aumentar a explicabilidade do modelo. Dessa forma, a rede proposta busca capturar a dinâmica temporal de atrofia cortical e outras alterações estruturais, evitando a dependência exclusiva de exames transversais e reduzindo a complexidade de abordagens longitudinais anteriores.

Os dados da base ADNI (ADNI-1/2/GO) com 235 pacientes, 487 pares longitudinais de RM ponderadas em T1, com intervalo de 1 ano entre T1 e T2. Os rótulos CCLe e CCLp utilizam o ponto de corte de 36 meses para conversão clínica. O pré-processamento incluiu correções de Gradwarp e B1, normalização N3, extração do encéfalo, padronização para MNI152 e correção não deformável. AFSNet adota DenseNet-121 como referência, acoplando LLFS após *DenseBlocks*, TFI para fusão temporal e DTAM para recalibrar canais por atenção. Os autores utilizaram o *framework* PyTorch, e treinamento com o otimizador Adam de 100 épocas, 8 amostras por *batch*, entropia cruzada (binária) como função de perda e validação cruzada em 5-*folds*. Nos resultados, o AFSNet supera ResNet-3D e VGG-3D sob as mesmas condições, com $0,7874 \pm 0,03$ de acurácia, $0,7958 \pm 0,03$ de F1-score e AUC de $0,7903 \pm 0,03$. O estudo de ablação identificou ganhos incrementais

de TFI, LLFS e DTAM sobre a DenseNet base. Os mapas de calor da Layer-CAM destacaram envolvimento crescente de regiões temporais e hipocampus no grupo CCLp.

O estudo de Agostinho, Simões e Castelo-Branco (2024) utilizou uma abordagem multimodal que integra dados de RM, PET, e imagem por tensor de difusão (DTI). Os dados foram obtidos da base ADNI, da qual foram selecionados 486 indivíduos com CCL, distribuídos entre os grupos CCLe e CCLp, utilizando o critério de 36 meses para seleção. Para o processamento das imagens de RM, os autores utilizaram o corregristo de um atlas de referência com 64 ROIs e segmentação para calcular o volume absoluto da GM e WM, que foram normalizados pelo ICV de cada indivíduo.

Modelos de SVMs foram construídos para cada modalidade individualmente e, posteriormente, combinados através de uma estratégia de fusão ponderada para avaliar a complementaridade entre elas. Os autores utilizaram um processo de validação que envolveu validação cruzada de 10 etapas. Para a modalidade de RM, o melhor desempenho foi obtido com os dados de GM, alcançando uma acurácia balanceada de $72,75 \pm 0,78$, F1-score de $0,77 \pm 0,01$ e AUC de $0,80 \pm 0,01$. A combinação da modalidade de RM com ambas as modalidades de PET (FDG e AV45) resultou em uma BACC de $78,79 \pm 0,84$, F1-score de $0,84 \pm 0,01$ e AUC de $0,88 \pm 0,01$. Por fim, na combinação de todas as quatro modalidades, o modelo obteve BACC de $71,84 \pm 0,77$, F1-score de $0,81 \pm 0,01$ e AUC de $0,87 \pm 0,01$. Dessa forma, a combinação de diferentes modalidades de imagens apresentou os melhores resultados, o que sugere que a integração de informações sobre alterações estruturais e metabólicas possui valor complementar na predição da conversão de CCL para DA.

3.3.2 Discussão

A literatura revisada sobre a progressão do CCL para a DA, apresentada na Tabela 6, revela uma evolução em direção a modelos computacionais com um forte predomínio de técnicas de aprendizado profundo. Nesse sentido, os estudos demonstram que a integração de múltiplas fontes de dados melhora o poder preditivo dos modelos. Essa abordagem multimodal é empregada de duas formas principais: (i) pela fusão de dados de imagem com informações de outras modalidades, como testes neuropsicológicos e biomarcadores de PET e DTI; e (ii) pela análise de dados longitudinais, que capturam a dinâmica temporal das alterações cerebrais. Os trabalhos de Luo et al. (2024) e Agostinho, Simões e Castelo-Branco (2024) exemplificam a primeira abordagem, mostrando ganhos significativos de acurácia ao combinar RM com resultados de testes neuropsicológicos ou com dados de PET, respectivamente. Já o estudo de Liu et al. (2025) foca na segunda abordagem, desenvolvendo uma arquitetura específica para modelar a progressão entre dois pontos no tempo, destacando a importância das mudanças estruturais ao longo do tempo.

Ao considerar os modelos utilizados, observa-se a ampla adoção de mecanismos de atenção nas arquiteturas das redes neurais. Estudos como os de Luo et al. (2023), Luo

et al. (2024), Rahim et al. (2025) e Liu et al. (2025) empregam diferentes mecanismos de atenção para guiar o foco do modelo às características e regiões cerebrais mais discriminativas, otimizando o processo de aprendizado. Somado a isso, a interpretabilidade dos modelos caixa-preta é uma preocupação significativa, especialmente no contexto clínico, com vários autores utilizando técnicas de visualização como CAMs para identificar as áreas anatômicas que mais influenciaram a predição. Essa busca por explicabilidade é fundamental para a futura validação e aceitação clínica desses modelos.

Apesar de resultados promissores, com alguns modelos atingindo acurácias superiores a 90%, a revisão sistemática de Wang et al. (2024) aponta para fragilidades metodológicas significativas na área, como a falta de validação externa, alto risco de viés e a falta de transparência, impedem que os modelos atuais sejam recomendados para uso clínico. Dessa forma, enquanto as pesquisas avançam na criação de arquiteturas complexas e multimodais que confirmam a importância de analisar múltiplas regiões cerebrais, como também aponta a revisão de Hassen et al. (2024), há ainda o desafio de aumentar o rigor metodológico, focar em validações independentes e garantir a reprodutibilidade para que esses avanços se traduzam em ferramentas confiáveis para aplicações clínicas.

Tabela 6 – Comparativo dos Trabalhos Correlatos em Predição da Progressão do CCL para a DA.

Autor(es)	Base de Dados	Metodologia Principal	Resultados Principais
Muhammed Niyas & Thiyagarajan (2023)	Revisão Sistemática (42 estudos)	Análise de métodos de aprendizado de máquina em dados longitudinais, com foco em abordagens multimodais.	Melhor desempenho encontrado foi de 96% de acurácia com dados multimodais.
Luo et al. (2023)	ADNI	CAAT (aprendizado profundo) com Mapas de Ativação de Classe (CAMs) para interpretabilidade.	Acurácia de 74,61% e AUC de 0,78. Superou uma 3D-CNN de controle, dispensando a segmentação de ROIs.
Luo et al. (2024)	ADNI	Mecanismo de dupla atenção para fundir características da imagem de RM com metadados neurocognitivos.	Acurácia de 81,34% e AUC de 0,874. A fusão de dados melhorou significativamente o desempenho.
Wang et al. (2024)	Revisão Sistemática (16 estudos)	Meta-análise de modelos preditivos para o risco de demência em idosos com CCL.	AUC combinada de 0,86. Modelos de aprendizado de máquina (AUC 0,89) superaram os tradicionais (AUC 0,77).
Agostinho et al. (2024)	ADNI	Abordagem multimodal (RM, PET, DTI) com SVMs e estratégia de fusão ponderada.	Melhor resultado com a fusão de RM e PET (AUC 0,88), demonstrando complementaridade entre modalidades.
Rahim et al. (2025)	ADNI	Ensemble de CNNs (2D) com múltiplas visualizações (axial, coronal, sagital) e mecanismo de atenção.	Acurácia de 91,77% e AUC de 92,87%. Mapas de atenção destacaram o hipocampo e córtex entorrinal.
Liu et al. (2025)	ADNI	AFSNet: CNN 3D com atenção para dados de RM longitudinais, focando na dinâmica temporal da atrofia.	Acurácia de 78,74% e AUC de 0,7903. Capturou a progressão entre dois pontos no tempo.

Fonte: elaborado pelo autor.

3.4 Considerações finais

Nesta revisão bibliográfica, foram exploradas três áreas de estudo para a análise computacional do CCL e DA. Primeiramente, constatou-se que as assimetrias inter-hemisféricas são biomarcadores robustos da neurodegeneração, com padrões específicos em regiões como o córtex entorrinal e o hipocampo, que se mostram capazes de diferenciar entre indivíduos CCLe e CCLp. Em seguida, a análise de detectores de pontos salientes 3D demonstrou que essas técnicas são eficazes em identificar locais de alta informação estrutural no encéfalo. Nesse sentido, pontos salientes podem servir como base para diferentes tarefas, desde o corregristo de imagens até a extração de características para modelos de classificação, superando abordagens que dependem exclusivamente de segmentações volumétricas. Finalmente, a terceira área de pesquisa abordou os modelos preditivos para a conversão de CCL para DA, evidenciando uma tendência para o uso de aprendizado profundo e a integração de dados multimodais e longitudinais. Com isso, a combinação de informações de diferentes modalidades, como testes neuropsicológicos e exames de PET, pode melhorar o desempenho dos modelos.

Contudo, apesar dos resultados promissores e do contínuo aprimoramento dos modelos de aprendizado profundo, persiste uma lacuna entre o desempenho obtido em ambientes de pesquisa e a aplicabilidade clínica. A falta de validação externa, o risco de viés e a necessidade de maior transparência metodológica são desafios críticos que precisam ser superados. Nesse contexto, o desenvolvimento de novas abordagens que integrem as frentes de estudo discutidas — como a utilização de pontos salientes para guiar a análise de assimetrias inter-hemisféricas e a aplicação de arquiteturas de aprendizado profundo para prever a conversão de CCL para DA — tem o potencial de avançar o estado da arte, contribuindo para um melhor entendimento da progressão da doença.

Capítulo 4

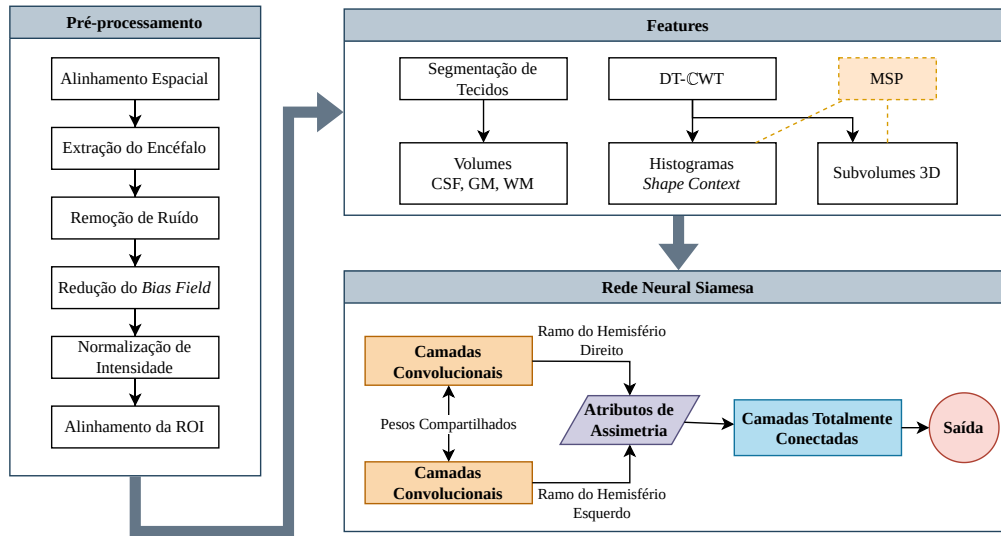
Materiais e Métodos

4.1 Enquadramento

Neste estudo, busca-se desenvolver uma técnica computacional automática que mapeie assimetrias estruturais inter-hemisféricas em imagens de RM. O objetivo é aprimorar a compreensão do processo de neurodegeneração na DA e, por conseguinte, contribuir para o seu diagnóstico. Para alcançar esse objetivo, foram empregadas técnicas de pré-processamento para garantir resultados consistentes, extração de características volumétricas via segmentação de tecidos, forma e textura via DT-CWT. A extração de atributos foi realizada na região hipocampal em cada um dos hemisférios cerebrais. Finalmente, esses atributos foram utilizados para o treinamento de modelos de aprendizado de máquina em tarefas de classificação binária nas populações CN, CCL e DA.

A Figura 9 apresenta um esquemático da abordagem desenvolvida para a detecção e análise de assimetrias inter-hemisféricas das principais estruturas anatômicas envolvidas na DA. Os elementos necessários para a realização desta tarefa incluem: (a) uma imagem de RM de referência; (b) um atlas contendo um conjunto de máscaras, alinhadas à imagem de referência, das estruturas anatômicas nas quais se planeja calcular a assimetria em cada hemisfério; (c) um banco de imagens alinhadas à imagem de referência e pré-processadas conforme a necessidade dos detectores e descritores utilizados; (d) um detector de pontos salientes 3D; (e) uma técnica para extração de características nos pontos detectados; (f) uma métrica para quantificar a similaridade entre descritores; (g) uma métrica para quantificar a assimetria entre os atributos extraídos de cada hemisfério.

Figura 9 – Fluxograma das principais etapas deste projeto de pesquisa.



4.2 Base de dados

Neste trabalho, são utilizadas imagens volumétricas de RM e modelos geométricos obtidos de três bases de domínio público, conforme descrito nas subseções a seguir.

4.2.1 ADNI

O projeto Alzheimer's Disease Neuroimaging Initiative (ADNI)¹ é uma iniciativa público-privada projetada para desenvolver biomarcadores em imagens biomédicas, genéticos e bioquímicos para a detecção precoce e acompanhamento da DA. Nesse estudo multicêntrico longitudinal, todos os dados gerados pelos núcleos de pesquisa são inseridos no repositório de dados hospedado pelo *Laboratory of Neuroimaging* (LONI) na *University of Southern California*. A base é composta por dados genéticos (*genome-wide association studies* e *whole genome sequencing*), imagens estruturais e funcionais de RM, imagens obtidas com tomografia por emissão de pósitrons e informações bioquímicas.

As imagens estruturais de RM foram adquiridas entre 2005 e 2018, obtidas pela sequência *Magnetization Prepared Rapid Gradient Echo* (MPRAGE) em *scanners* de 1.5T e 3T. Essas imagens foram adquiridas em estudos longitudinais de 559 indivíduos pertencentes às classes CN, CCL e DA, totalizando 2350 imagens. Estes exames estão distribuídos em quatro grupos diagnósticos: 698 imagens de indivíduos CN, 582 de indivíduos com CCL, 653 de indivíduos CCLp, e 417 de indivíduos com diagnóstico de DA. O grupo CCLp compreende 136 indivíduos que converteram para a DA dentro de um período de 36 meses, enquanto o grupo CCL (130 indivíduos) permaneceu clinicamente estável no mesmo intervalo. Neste estudo, foram utilizadas apenas as imagens estruturais T1-w. Os

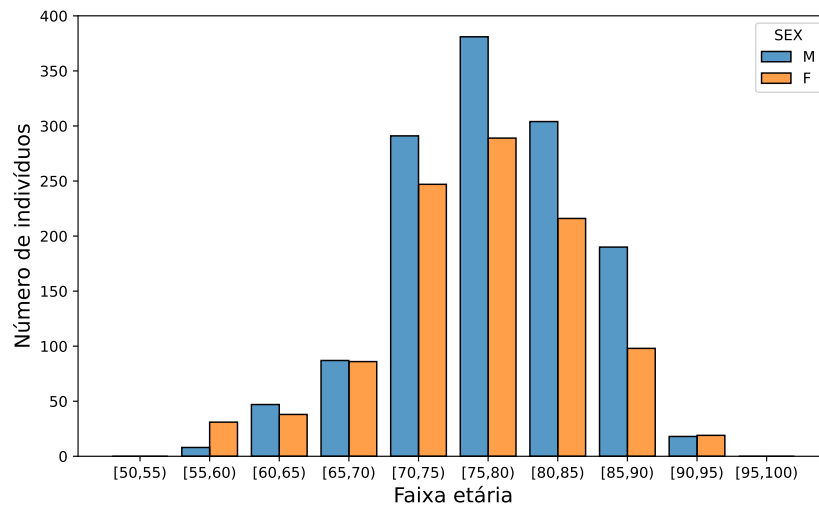
¹ Disponível em: <https://adni.loni.usc.edu/> Acesso em: 01 dez, 2023.

Tabela 7 – Informações demográficas das imagens de RM da base ADNI.

	CN	CCL	DA	Total
No. indivíduos	159	264	136	559
Gênero (F : M)	82 : 77	99 : 165	63 : 73	244 : 315
Idade	$77,66 \pm 5,42$	$76,42 \pm 7,39$	$76,41 \pm 7,45$	$74,4 \pm 7,6$

Fonte: elaborada pelo autor.

Figura 10 – Informações demográficas das imagens da base ADNI.



Fonte: elaborada pelo autor.

dados demográficos das imagens utilizadas são apresentados na Tabela 7 e Figura 10.

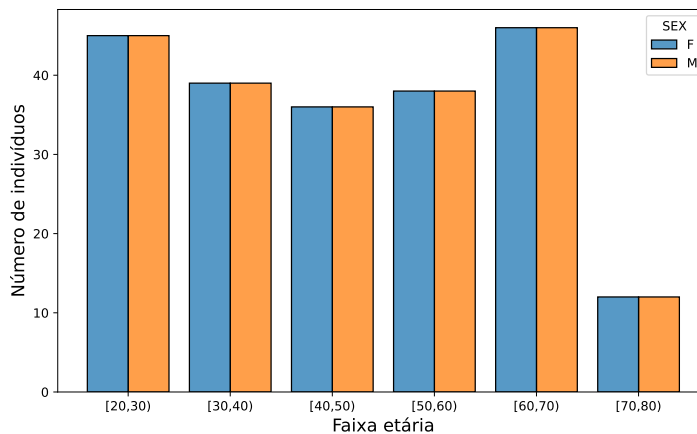
Foram selecionados, aleatoriamente, 40 indivíduos (20 CN e 20 DA) para etapas de otimização de atributos. Totalizando 67 imagens da classe CN e 62 da classe DA, que foram excluídas dos conjuntos de testes. Dessa forma, o conjunto de imagens de teste foi composto por 129 indivíduos CN (631 imagens) e 116 indivíduos da classe DA (355 imagens).

4.2.2 IXI

A base de imagens *Information eXtraction from Images* (IXI)² possui 532 imagens de RM obtidas de indivíduos saudáveis em três hospitais de Londres, sendo eles *Hammersmith Hospital* (HH), *Guy's Hospital* (GH) e *Institute of Psychiatry* (IoP). Para cada uma dessas imagens, foram disponibilizadas versões nas modalidades T1-w, T2-w, PD-w, MRA e DTI. Neste trabalho, as imagens T1-w dessa base são utilizadas para a construção do atlas probabilístico de pontos salientes. Os dados demográficos são apresentados na Figura 11.

² Disponível em: <https://brain-development.org/ixi-dataset/> Acesso em: 01 dez, 2023.

Figura 11 – Informações demográficas das imagens de base IXI.



Fonte: elaborado pelo autor.

4.2.3 SPL/NAC

O projeto *Open Antomy Project* ³ é um projeto que visa disponibilizar atlas anatômicos por meio de dados abertos, desenvolvimento colaborativo e distribuição gratuita de conhecimento médico. O *Neuroimaging Analysis Center* (NAC) em colaboração com o *Surgical Planning Laboratory* (SPL) disponibilizaram o atlas anatômico SPL/NAC ⁴ de um voluntário masculino saudável de 42 anos. As imagens foram adquiridas com um *scanner* de RM Siemens de 3T e as malhas triangulares foram geradas automaticamente com o software *FreeSurfer*, ajustadas e rotuladas manualmente por um especialista.

Nesse projeto, uma imagem T1-w, com resolução isotrópica de 1 mm e de tamanho $256 \times 256 \times 256$, foi utilizada como referência. Além disso foram utilizados dois modelos geométricos (malhas) da segmentação do hipocampo, direito e esquerdo, para selecionar a região de interesse que é analisada.

4.3 Pré-processamento das imagens

A etapa inicial da metodologia consistiu em um *pipeline* de pré-processamento projetado para padronizar as imagens de RM clínicas e melhorar sua qualidade para análises subsequentes. Este *pipeline* compreende seis etapas sequenciais: alinhamento espacial, remoção do crânio (*skull stripping*), redução de ruído (*denoising*), correção do *bias field*, normalização de intensidade e alinhamento dos rótulos do atlas. Cada etapa é descrita a seguir.

³ Disponível em: <https://www.openanatomy.org/about.html> Acesso em 01 dez, 2023.

⁴ Disponível em: <https://www.openanatomy.org/atlas-pages/atlas-spl-nac-brain.html> Acesso em 01 dez, 2023.

4.3.1 Alinhamento espacial

Para estabelecer um sistema de coordenadas anatômico comum entre todos os sujeitos, cada imagem T1-w foi alinhada ao atlas de referência da NAC utilizando um procedimento de duas etapas, implementado através do pacote de software Elastix, um framework de código aberto para registro de imagens médicas (KLEIN et al., 2009; SHAMONIN et al., 2014). Inicialmente, as imagens foram alinhadas de forma aproximada, combinando seu centro de massa com o do atlas de referência. Em seguida, o corregristo afim global, que corrige a variabilidade espacial entre indivíduos e preserva as estruturas anatômicas e a simetria inter-hemisférica. A transformação não envolveu deformações não lineares, evitando assim distorções que poderiam enviesar as análises de assimetria subsequentes.

Este alinhamento inicial também desempenha um papel crítico em melhorar o desempenho da etapa de remoção do crânio. Ao reduzir a variabilidade anatômica e ao colocar as principais estruturas em alinhamento aproximado com o espaço do atlas, ele permite uma identificação e remoção mais precisa dos tecidos não pertencentes ao encéfalo. Para o corregristo afim, foi utilizado o conjunto de parâmetros "Par0010", que é otimizado para imagens de RM 3D do encéfalo e está disponível através do ElastixModelZoo⁵.

4.3.2 Extração do encéfalo (*Skull Stripping*)

Após a normalização espacial, os tecidos não pertencentes ao encéfalo foram removidos utilizando o Synth-Strip, um algoritmo de remoção de crânio baseado em aprendizado profundo, descrito na Seção 2.5.5. O SynthStrip sintetiza *priors* anatômicos específicos do sujeito para gerar máscaras precisas do encéfalo.

4.3.3 Redução de ruído (*denoising*)

A técnica NLM, descrita na Seção 2.5.1 foi utilizada para realizar a redução de ruído. O algoritmo foi configurado para considerar a distribuição Rician do ruído, que é característica de dados de RM reconstruídos em magnitude. Neste trabalho, o filtro foi aplicado com um raio de *patch* de 1 *voxel* e um raio de busca de 3 *voxels*. O fator de redução (*shrink factor*) foi definido como 1 para manter a resolução espacial original das imagens. Esses parâmetros foram escolhidos para alcançar um equilíbrio entre a supressão eficaz de ruído e a preservação de detalhes estruturais finos, especialmente na região hipocampal. A otimização desses parâmetros foi realizada por meio de uma busca em grade em um subconjunto de 8 imagens (4 CN e 4 DA), amostradas aleatoriamente a partir do conjunto de 40 indivíduos reservado para testes no banco de dados ADNI. O critério de otimização priorizou a maximização da quantidade de pontos salientes detectados, seguida de uma inspeção visual das imagens, em ordem decrescente de pontos salientes, para identificar

⁵ <<https://github.com/SuperElastix/ElastixModelZoo>>

a combinação de parâmetros que melhor suprimiu o ruído e preservou a estrutura do hipocampo.

4.3.4 Correção do *bias field*

Para corrigir as inomogeneidades de intensidade de baixa frequência comumente presentes em imagens de RM, empregamos o algoritmo N4ITK, descrito na Seção 2.5.2. O algoritmo foi configurado com um fator de redução de 3 e um cronograma multirresolução de três níveis com 20 iterações por nível. A convergência foi definida com base em um limiar de mudança relativa de parâmetro de 1×10^{-3} . Adicionalmente, o reescalamento da intensidade dentro da máscara cerebral foi ativado para melhorar a estabilidade numérica e a normalização entre os sujeitos.

4.3.5 Normalização de intensidade

A técnica de normalização baseada em FCM, descrita na Seção 2.5.3, utiliza a distribuição de intensidade de WM como referência para reescalonar as intensidades da imagem, harmonizando o brilho e o contraste em todo o conjunto de dados de maneira automática.

4.3.6 Identificação da região de interesse anatômica

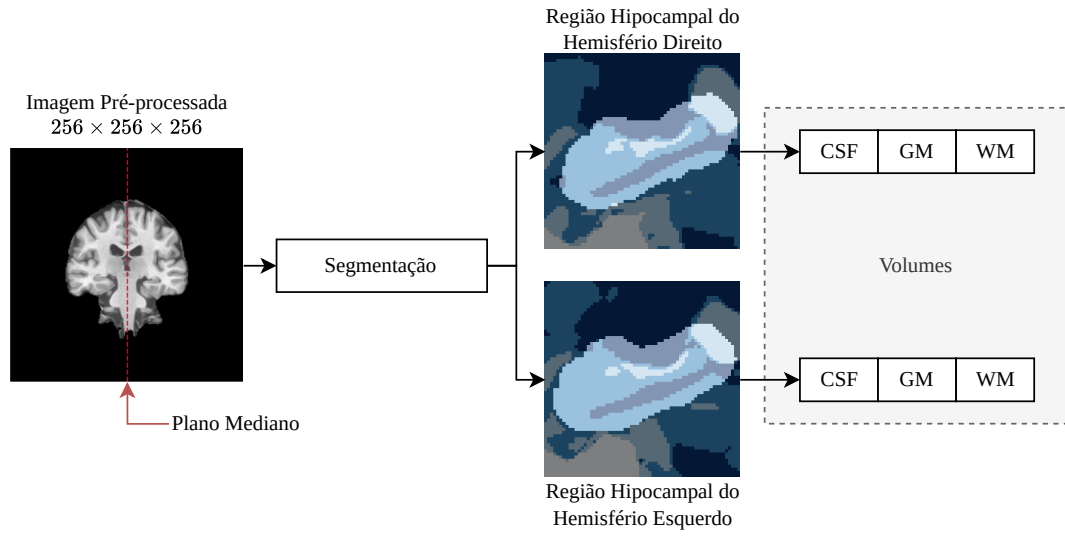
Em contraste com o alinhamento inicial da imagem, Seção 4.3.1, que empregou apenas transformações rígidas e afins (não deformáveis) para preservar a geometria anatômica e permitir uma normalização global, esta etapa aplica um registro deformável para definir com precisão as regiões de interesse específicas de cada sujeito para a análise de assimetria. Especificamente, o atlas SPL/NAC, que inclui segmentações anatômicas rotuladas, foi deformado não linearmente para se ajustar à imagem pré-processada de cada sujeito usando uma transformação *B-spline*. Neste caso, a imagem do sujeito serviu como referência fixa para garantir que os detalhes anatômicos permanecessem sem distorção, enquanto o atlas foi tratado como a imagem móvel. Essa abordagem permite a localização precisa de estruturas anatômicas no cérebro de cada indivíduo, adaptando as máscaras do atlas à morfologia específica do sujeito. Como tal, este registro deformável suporta eficazmente análises subsequentes, como a parcellação cerebral e o cálculo de assimetria baseado em ROI, sem introduzir artefatos espaciais no espaço nativo da imagem do indivíduo.

4.4 Segmentação e quantificação dos tecidos cerebrais

As características volumétricas foram extraídas por meio da segmentação do cérebro em CSF, GM e WM utilizando o algoritmo Atropos (AVANTS et al., 2011), disponível no

biblioteca ANTs⁶. Neste trabalho, o programa foi executado com as configurações $K = 3$, fator de suavização com raio de $1 \times 1 \times 1$ e, $c = 10$ EM iterações. Esses parâmetros foram seleccionados considerando o equilíbrio entre qualidade da segmentação e tempo de processamento de cada imagem. Após a execução, o resultado é uma imagem com valores inteiros que codificam o plano de fundo (*background*), CSF, GM, e WM, em que calcula-se o volume de cada tecido cerebral dentro da ROI via contagem de *voxels* pertencentes à cada tecido. A Figura 4.4 apresenta um esquemático dos atributos extraídos.

Figura 12 – Principais etapas empregadas na quantificação dos tecidos cerebrais



Fonte: elaborado pelo autor.

4.5 Detecção de pontos salientes com a DT-CWT

Fauqueur, Kingsbury e Anderson (2006) propuseram uma metodologia que utiliza a transformada DT-CWT 2D para decompor uma imagem bidimensional e gerar um mapa de energia acumulada, definido como:

$$E(C) = \alpha^s \left(\prod_{b=1}^6 \rho_b \right)^\beta, \quad (59)$$

em que α e β são parâmetros ajustáveis que ponderam o acúmulo de cada escala s , e $C = \{\rho_1 e^{i\theta_1}, \dots, \rho_6 e^{i\theta_6}\}$ é o conjunto que contém a resposta direcional em seis orientações: $15^\circ, 45^\circ, 75^\circ, 105^\circ, 135^\circ, 160^\circ$ (FAUQUEUR; KINGSBURY; ANDERSON, 2006). O mapa de energia $E(C)$ é calculado para diferentes escalas via reamostragem (*downsampling*) da imagem original, reduzindo-a por fatores de dois, sendo $s = 1, \dots, m$ a quantidade de escalas, então o tamanho da imagem na escala s é dado por $\frac{N_x}{2^s} \times \frac{N_y}{2^s}$, em que N_x e

⁶ <<https://github.com/ANTsX/ANTsPy>>

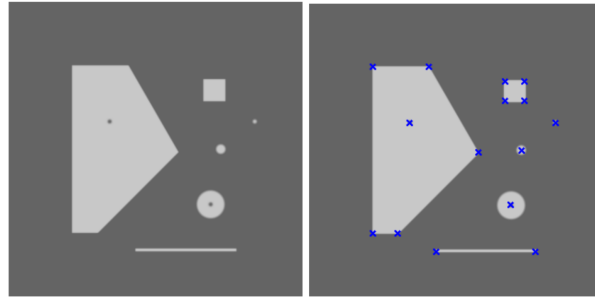
N_y representam as dimensões da imagem original (FAUQUEUR; KINGSBURY; ANDERSON, 2006).

Sendo M_s o mapa de energia $E(C)$ obtido na escala s , então o mapa de energia acumulada é definido como:

$$A = \sum_{s=1}^m f_s(M_s), \quad (60)$$

sendo $f_s(M_s)$ um *kernel* Gaussiano 2D que realiza a interpolação da imagem (*upsampling*) para aumentar a resolução espacial por um fator de 2^s (FAUQUEUR; KINGSBURY; ANDERSON, 2006). O intuito dessa técnica é gerar uma imagem de tamanho $w \times h$, cujos valores de intensidade representam uma mistura de gaussianas em que cada pico representa um ponto saliente que pode ser extraído calculando-se o máximo local em uma vizinhança $n \times n$ (FAUQUEUR; KINGSBURY; ANDERSON, 2006). Uma representação visual do mapa de energia gerado para a imagem da Figura 4.5(à esquerda) é apresentada na Figura 4.5.

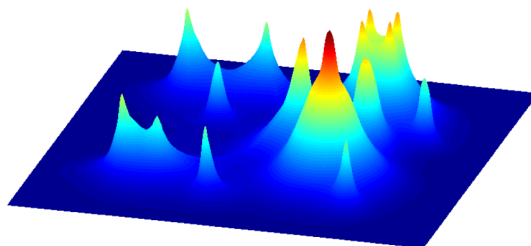
Figura 13 – Imagem sintética ilustrativa do detector DT-CWT. Esquerda: imagem original. Direita: pontos detectados e sobrepostos na imagem original.



Fonte: Fauqueur, Kingsbury e Anderson (2006).

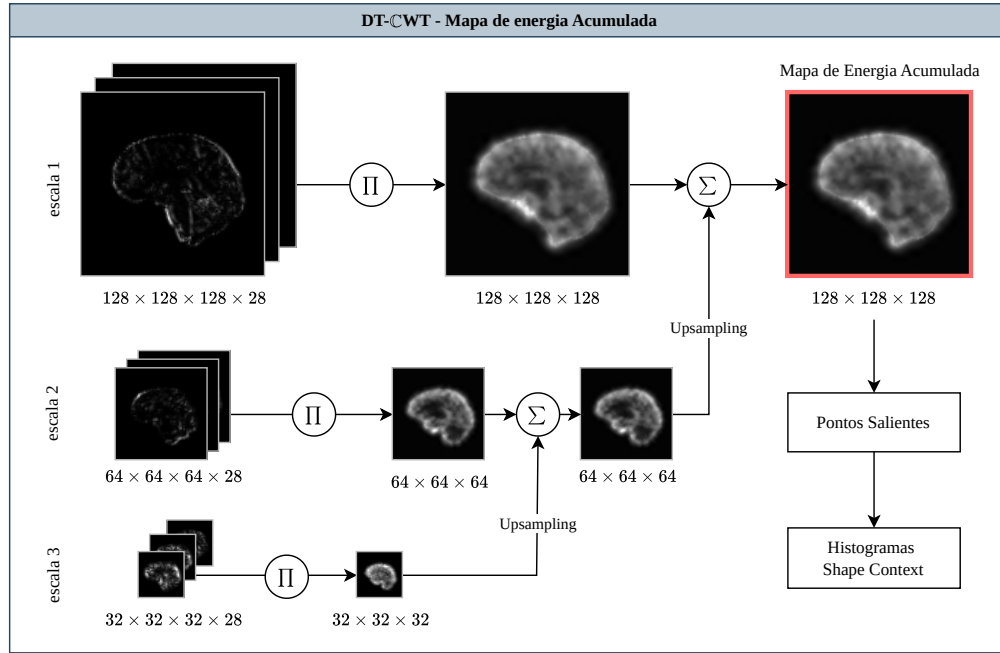
Para sinais tridimensionais (3D), Souza, Poloni e Ferrari (2020) utilizaram uma abordagem semelhante, na medida que os pontos salientes são detectados no mapa de energia acumulada. Dessa forma, dada uma imagem de tamanho $N_x \times N_y \times N_z$, o sinal é decomposto pelos coeficientes da DT-CWT 3D em m escalas, em que $\frac{N_x}{2^s} \times \frac{N_y}{2^s} \times \frac{N_z}{2^s}$ o tamanho da imagem na escala s (SOUZA; POLONI; FERRARI, 2020). Em seguida, o mapa de

Figura 14 – Mapa de energia acumulada A para sinais 2D.



Fonte: Fauqueur, Kingsbury e Anderson (2006).

Figura 15 – Fluxograma das principais técnicas necessárias para construção do mapa de energia acumulada.



Fonte: elaborado pelo autor.

energia acumulada é definido como:

$$E(C) = \alpha \left(\prod_{b=1}^{28} \rho_b \right)^\beta, \quad (61)$$

em que $C = \{\rho_1 e^{i\theta_1}, \dots, \rho_{28} e^{i\theta_{28}}\}$ é o conjunto de respostas direcionais geradas pela transformada (SOUZA; POLONI; FERRARI, 2020).

O somatório definido na equação 60 é utilizado para gerar o mapa de energia acumulada, assim como no detector 2D. Finalmente, as posições de máximo locais no mapa de energia resultante A , após a eliminação dos *voxels* com intensidade inferior ao limiar T_E , são utilizadas para definir as posições dos pontos salientes na imagem original (SOUZA; POLONI; FERRARI, 2020). Antes da avaliação com o limiar T_E , aplica-se uma normalização *MinMax* para garantir que os valores em A estejam limitados ao intervalo $[0, 1]$. Na prática, a normalização em todos os *voxels* de A pode não ser desejada, uma vez que um único *voxel* com intensidade altamente elevada é capaz de excluir a detecção de pontos em todas as demais regiões da imagem. Dessa forma, uma máscara pode ser utilizada para restringir a normalização à um único objeto ou região em A .

Neste trabalho, foi utilizada uma versão modificada do mapa de energia acumulada, em que, não é aplicada a última etapa de *upsampling*, i.e., o mapa de energia acumulada gerado possui o mesmo tamanho $\frac{N_x}{2} \times \frac{N_y}{2} \times \frac{N_z}{2}$. A Figura 4.5 apresenta um esquemático das etapas descritas na construção do mapa de energia acumulada.

4.6 Atlas probabilístico de pontos salientes

Devido à variações anatômicas entre pacientes e na qualidade das imagens, a criação de um atlas probabilístico de pontos salientes 3D visa avaliar a estabilidade dos pontos detectados. Assim, a construção do atlas permite caracterizar a normalidade dos pontos, considerando apenas aqueles com alta ocorrência na população normal. Para isso, utilizando imagens da população CN, o algoritmo proposto por Villa Pinto e Ferrari (2016) calcula uma estimativa para diversas funções de densidade de probabilidade locais que expressam a ocorrência dos pontos nessas imagens. Após essa estimativa, os pontos estáveis são determinados considerando a localização espacial dos picos de máximo locais no atlas probabilístico resultante.

A função de densidade de probabilidade é estimada usando o modelo não-paramétrico *Manifold Parzen Windows* (MPW), em que $p(\mathbf{x})$ representa um conjunto de ℓ amostras d -dimensional de misturas de Gaussianas, $G_{\mu, \Sigma}$, e expressa a ocorrência dos pontos na população, é definida como

$$p(\mathbf{x}) = \frac{1}{\ell} \sum_{i=1}^{\ell} G_{\mu=\mathbf{x}_i, \Sigma=\Sigma_i}(\mathbf{x}) \quad (62)$$

$$= \frac{1}{\ell} \sum_{i=1}^{\ell} \frac{1}{\sqrt{(2\pi)^d \det(\Sigma_i)}} \exp \left[-\frac{1}{2} \Delta(\mathbf{x}) \right], \quad (63)$$

sendo $\Delta(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_i)^T \Sigma_i^{-1} (\mathbf{x} - \mathbf{x}_i)$ a distância quadrada de *Mahalanobis* e $\Sigma_i = \Sigma_{\mathcal{K}_i} + \sigma^2 \mathbf{I}_3$ representa a matriz de covariância do i -ésimo componente Gaussiano. Com isso, para que esses componentes sejam dados em função das direções principais das amostras que estão na vizinhança de $\mathbf{x}_i$, o componente $\Sigma_{\mathcal{K}_i}$ da matriz de covariância de Σ_i é dado por

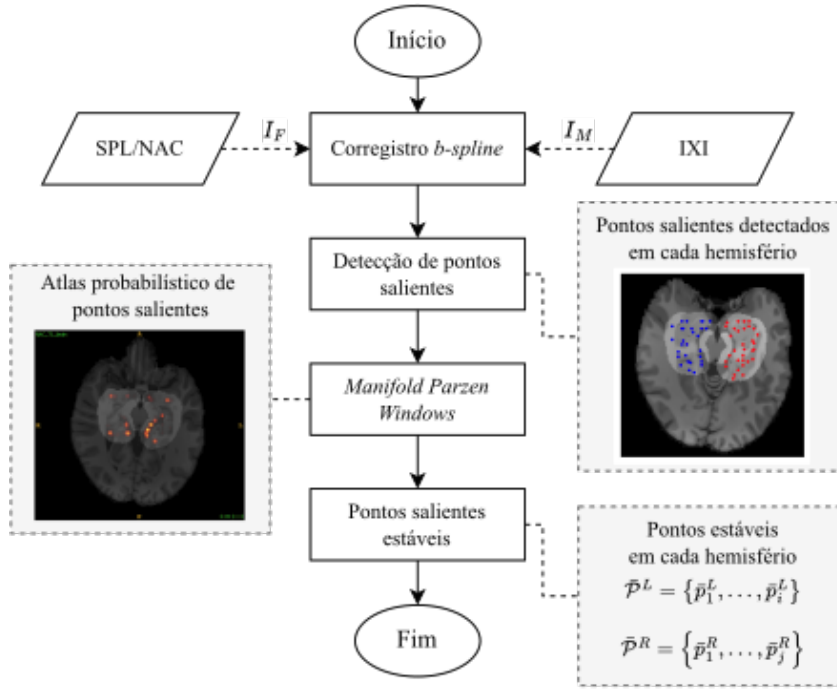
$$\Sigma_{\mathcal{K}_i} = \frac{\sum_{j=1 \dots \ell, j \neq i} \mathcal{K}(\mathbf{x}_j, \mathbf{x}_i) (\mathbf{x}_j - \mathbf{x}_i)^T (\mathbf{x}_j - \mathbf{x}_i)}{\sum_{j=1 \dots \ell, j \neq i} \mathcal{K}(\mathbf{x}_j, \mathbf{x}_i)}, \quad (64)$$

em que $\mathcal{K}(\mathbf{x}, \mathbf{x}_i)$ é uma função de vizinhança que associa uma ponderação a cada amostra $\mathbf{x}$ na vizinhança de $\mathbf{x}_i$, podendo limitar essa vizinhança às k_{cov} amostras mais próximas (VINCENT; BENGIO, 2002).

Após o pré-processamento das imagens, as bases de dados IXI e NAC/SPL são utilizadas para construir um atlas probabilístico de pontos salientes, do qual foram extraídos os pontos estáveis na população CN. A detecção desses pontos é feita seguindo três etapas: (i) detecção dos pontos salientes nas imagens alinhadas, (ii) construção do atlas probabilístico de pontos salientes, (iii) posicionamento dos pontos sob a imagem de referência.. A Figura 4.6 apresenta um esquemático dessas etapas.

Em cada imagem clínica da base IXI, realiza-se a detecção dos pontos saliente com a técnica DT-CWT 3D, como descrito na Sub-seção 4.5. Antes de realizar a localização das posições de máximo locais no mapa de energia acumulada, aplica-se uma normalização local delimitada pelas máscaras M_s^L e M_s^R . Em seguida, os pontos salientes são detectados

Figura 16 – Fluxograma da sequência de etapas utilizadas para a detecção dos pontos estáveis na população CN.



Fonte: elaborado pelo autor.

nesse mapa normalizado. Assim, para cada imagem clínica, os pontos salientes detectados em uma determinada estrutura s , são representados pelos conjuntos $\mathcal{P}^L = \{p_1^L, \dots, p_n^L\}$ e $\mathcal{P}^R = \{p_n^R, \dots, q_j^R\}$, em que p_n^L e p_n^R representam a posição do ponto saliente no hemisfério esquerdo e direito, respectivamente.

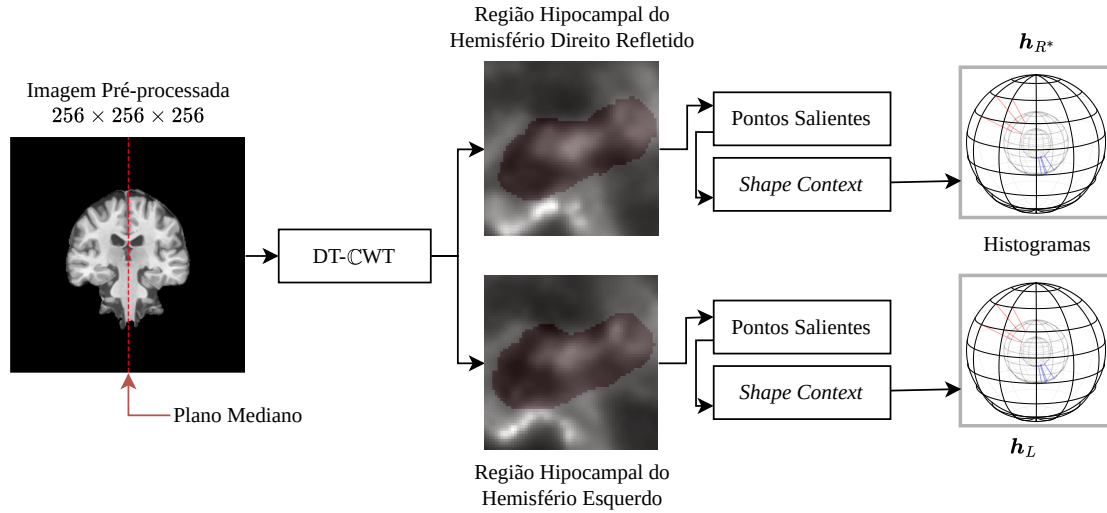
Embora o corregristo afim já tenha sido utilizado para alinhar as imagens da base IXI à imagem de referência SPL/NAC, o corregristo deformável *B-spline* é capaz de obter uma melhor sobreposição das imagens, uma vez que cada ponto de suporte realiza a transformação localmente na imagem móvel. Nesse sentido, os parâmetros obtidos e salvos por essa mesma transformação são aplicados novamente para mover os pontos da imagem clínica para as coordenadas da imagem de referência.

Após a detecção dos pontos em todas as imagens e estruturas, a técnica MPW, descrita na Seção 4.6, é usada para estimar a função densidade de probabilidade dos pontos salientes, resultando no atlas probabilístico de pontos salientes em cada hemisfério cerebral. Assim, os pontos estáveis são aqueles localizados nas regiões de máximo local do atlas probabilístico, representados pelos conjuntos $\bar{\mathcal{P}}^L = \{\bar{p}_1^L, \dots, \bar{p}_i^L\}$ e $\bar{\mathcal{P}}^R = \{\bar{p}_1^R, \dots, \bar{p}_j^R\}$ que correspondem, respectivamente, aos conjuntos de pontos estáveis dos hemisférios esquerdo e direito.

Apenas imagens de indivíduos com idade superior à 50 anos foram utilizadas na criação do atlas probabilístico de pontos salientes. Essa idade foi escolhida para que os pontos salientes estáveis fossem mais representativos dos indivíduos pertencentes à base ADNI.

4.7 Descritor *shape context*

Figura 17 – Principais etapas empregadas na extração dos descritores *shape context*.



Fonte: elaborado pelo autor.

4.7.1 Posicionamento do descritor *shape context*

A eficácia do descritor *shape context* depende fundamentalmente da escolha do seu ponto de referência, uma vez que o histograma log-esférico gerado é altamente sensível a essa localização. Uma referência bem posicionada em uma região anatômica relevante garante a extração de um vetor de características mais robusto e discriminativo.

Neste trabalho, foram exploradas e comparadas quatro estratégias distintas para definir o ponto de referência p_0 do descritor *shape context* dentro da região de interesse analisada.

4.7.1.1 Pontos salientes estáveis

Esta estratégia utilizou como referência os pontos salientes estáveis da população CN. A abordagem baseia-se no atlas probabilístico de pontos salientes, conforme descrito na Seção 5.3, utilizando imagens da base IXI. O processo consiste nas seguintes etapas:

1. **Mapeamento para a imagem clínica:** Os pontos estáveis, definidos no espaço do atlas, são mapeados para cada imagem clínica da base ADNI por meio da transformação deformável *B-spline*, obtida durante o corregristo.
2. **Otimização local:** Para refinar a localização e adaptá-la às particularidades anatômicas de cada indivíduo, realiza-se uma busca local por um pico de intensidade no mapa de energia acumulada da DT-CWT. Uma janela de busca é centralizada na

posição do ponto mapeado, e a coordenada de máximo local dentro dessa janela é definida como o ponto de referência final para o descritor.

Essa abordagem visa posicionar o descritor em locais de alta relevância estrutural, que são posições estáveis na população CN, permitindo uma análise mais precisa das deformações causadas pela doença. As correspondências inter-hemisféricas foram definidas manualmente na imagem de referência SPL/NAC.

4.7.1.2 Pontos salientes (por indivíduo)

Em contraste com a abordagem anterior, esta estratégia não utiliza o atlas probabilístico. Em vez disso, os pontos de referência são definidos pelos pontos salientes detectados diretamente na imagem de cada indivíduo.

Para cada ROI, o detector DT-CWT 3D é aplicado para gerar um mapa de energia local. Os máximos locais neste mapa são identificados como o conjunto de pontos salientes para aquele sujeito específico. O descritor *shape context* é então calculado para cada um desses pontos. A vantagem desta técnica é sua sensibilidade à morfologia única de cada cérebro, capturando características que podem ser específicas de um indivíduo ou de um estágio particular da doença.

As correspondências inter-hemisféricas foram definidas com a técnica descrita na Seção 2.8.1, sendo o descritor utilizado para obter o casamento, aquele formado pelo mapa de energia da segunda escala ($64 \times 64 \times 64$).

4.7.1.3 Centroide da máscara anatômica

Estratégia direta e global, utilizando um único ponto de referência para toda a estrutura anatômica. O ponto é definido como o centro de massa (centroide) da máscara binária da ROI.

A posição do centroide (p_0) é calculada pela média aritmética das coordenadas de todos os *voxels* (v_i) que compõem a máscara da estrutura. Essa abordagem garante um ponto de referência robusto a ruídos e a pequenas variações de forma.

4.7.1.4 Centroide e pontos salientes

Esta abordagem híbrida busca combinar a robustez global do centroide da ROI com a sensibilidade local dos pontos salientes. O ponto de referência final, p_0 , é definido via combinação linear de pontos salientes (c_p) detectados na máscara binária. O cálculo é definido pela seguinte equação:

$$p_0 = a \cdot c_p + b \cdot c_r, \quad (65)$$

em que, a e b são pesos que definem a contribuição de cada centroide, permitindo equilibrar a informação da forma geral da estrutura anatômica (via c_r) e a distribuição de suas características locais mais proeminentes (via c_p).

4.8 Classificador XGBoost

O modelo XGBoost, descrito na Seção 2.9, foi utilizado como classificador binário com os atributos de forma (*shape context*) e volumétricos. Esse modelo foi implementado para estabelecer um referencial de desempenho para avaliar, por comparação, a eficácia dos modelos mais complexos baseados em redes neurais. A otimização dos hiperparâmetros do XGBoost foi conduzida via busca em grade, explorando a combinação de atributos definidas pelos seguintes valores:

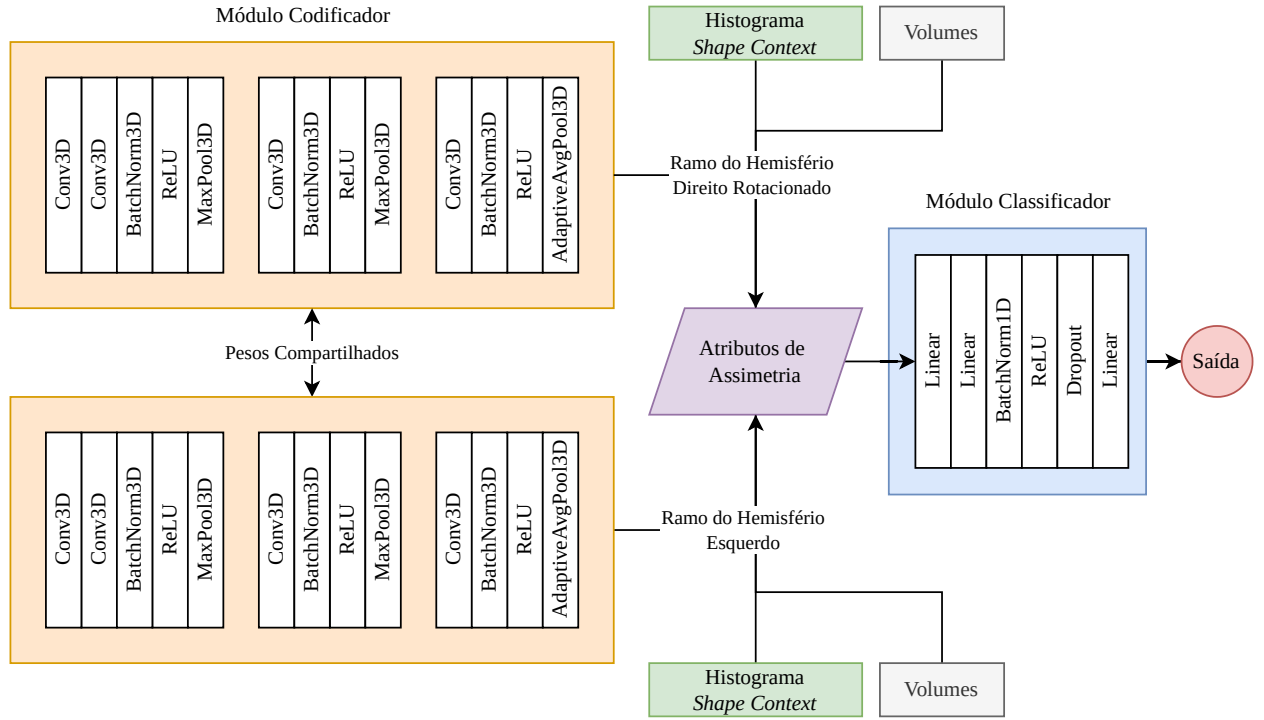
- ❑ `n_estimators`: {50, 100, 200}
- ❑ `max_depth`: {4, 5, 6, 7}
- ❑ `learning_rate`: {0.05, 0.1, 0.2, 0.3}
- ❑ `colsample_bytree`: {0.7, 0.8, 0.9, 1.0}
- ❑ `subsample`: {0.6, 0.7, 0.8, 0.9, 1.0}

4.9 Modelagem da neural Siamesa

Um dos principais desafios na aplicação de técnicas de aprendizado profundo para o diagnóstico da DA reside na disponibilidade limitada de dados anotados, especialmente quanto à distinção de alterações estruturais sutis em indivíduos com CCL. Essa limitação é exacerbada pelas pequenas diferenças anatômicas entre os grupos de indivíduos CN, com CCLe e com CCLp.

Para modelar explicitamente as assimetrias estruturais inter-hemisféricas, este trabalho propõe uma arquitetura de rede Siamesa, denominada Rede Neural Siamesa Convolutiva Siamesa (*Siamese Convolutional Neural Network* - SCNN), ilustrada na Figura 18. Esta arquitetura foi projetada para uma análise multimodal, integrando biomarcadores de diferentes naturezas. A rede processa, em paralelo, dados provenientes de regiões cerebrais homólogas dos hemisférios esquerdo e direito. Como passo de pré-processamento, são extraídos *patches* de $48 \times 48 \times 48$ *voxels* centrados no hipocampo de cada hemisfério a partir do mapa de energia acumulada da DT-CWT. Para garantir uma orientação consistente para a análise de assimetria, o *patch* correspondente ao hipocampo direito é refletido em relação ao plano sagital mediano.

Figura 18 – Arquitetura SCNN multimodal.



Fonte: elaborado pelo autor.

A base da arquitetura consiste em dois ramos de CNNs 3D, que são idênticos e compartilham os mesmos pesos. Cada ramo processa de forma independente os dados de um hemisfério, aprendendo a extrair representações de características de alto nível. A decisão de incorporar os coeficientes da DT-CWT diretamente como canais de entrada da rede se fundamenta no trabalho de (SINGH; KINGSBURY, 2017), que demonstrou que tal abordagem confere ao modelo uma vantagem inicial, resultando em ganhos de eficiência e aceleração da convergência durante o treinamento.

Após a extração de características pelos blocos convolucionais de cada ramo, os vetores de atributos de alto nível são concatenados com seus respectivos descritores de forma (*shape context*) e volume tecidual. Para codificar a assimetria de maneira explícita, os vetores de características combinados de cada hemisfério são fundidos por meio do cálculo da diferença absoluta, realizada elemento a elemento. Nesse sentido, a comparação inter-hemisférica, via diferença absoluta, segue o padrão estabelecidos por outros estudos apresentados na Seção 3.1. Este processo, gera um único vetor de assimetria, que representa de forma unificada as disparidades inter-hemisféricas em todas as modalidades analisadas.

Finalmente, na etapa de classificação, este vetor de assimetria multimodal é direcionado a uma sequência de camadas densas (*fully connected layers*), que são responsáveis

por mapear a representação de assimetria aprendida para um rótulo de diagnóstico clínico binário. A implementação da rede foi realizada utilizando o *framework* PyTorch⁷, e sua arquitetura detalhada é apresentada na Tabela 8.

Tabela 8 – Arquitetura detalhada da DT-CWT SCNN. A tabela descreve os blocos convolucionais do módulo codificador e as camadas do módulo classificador.

Nome da Camada	Parâmetros
Conv3D	$3 \times 3 \times 3$ kernels, stride=2, padding=1, 32 canais de saída
Conv3D	$3 \times 3 \times 3$ kernels, stride=2, padding=1, 64 canais de saída
BatchNorm3D	
ReLU	
MaxPool3D	$2 \times 2 \times 2$ kernels, stride=2
Conv3D	$3 \times 3 \times 3$ kernels, stride=2, padding=1, 128 canais de saída
BatchNorm3D	
ReLU	
MaxPool3D	kernel=2, stride=2
Conv3D	$3 \times 3 \times 3$ kernels, stride=2, padding=1, 256 canais de saída
BatchNorm3D	
ReLU	
AdaptiveAvgPool3D	$3 \times 3 \times 3$, 256 canais de saída
Linear	$(256 \times 3 \times 3 \times 3, 1024)$
Linear	$(1024, 512)$
BatchNorm1D	
ReLU	
Dropout (30 %)	
Linear	$(512, 1)$

Fonte: elaborado pelo autor.

4.9.1 Modelo de Referência

Com o intuito de quantificar a contribuição das características multimodais propostas, foi implementado um modelo de referência (*baseline model*). Este modelo consiste em uma rede Siamesa CNN que preserva a mesma arquitetura de ramos duplos e o mesmo bloco de classificação da rede principal. A diferença fundamental reside na sua entrada: nesta configuração, apenas os *patches* de imagem 3D de $64 \times 64 \times 64$ voxels, extraídos diretamente das imagens T1-w pré-processadas, foram fornecidos a cada ramo. Foram deliberadamente excluídos os coeficientes da DT-CWT, os descritores *shape context* e os volumes teciduais. Portanto, este modelo de referência funciona como um controle experimental, permitindo isolar e avaliar o impacto da integração das características de forma, volume e textura no desempenho diagnóstico da arquitetura completa.

⁷ <<https://pytorch.org/>>

4.10 Atributos inter-hemisféricos

Para obter uma representação discriminativa da assimetria cerebral, foi realizada a fusão de atributos de três modalidades distintas, extraídos dos hemisférios esquerdo e direito. Este processo combina descritores projetados manualmente (*handcrafted*) com características aprendidas por redes neurais profundas (*deep-learned features*), consolidando-os em um único vetor que quantifica explicitamente as diferenças inter-hemisféricas.

A fim de estabelecer um referencial espacial consistente para a extração de atributos, o hemisfério direito da imagem de RM de cada indivíduo foi espelhado em relação ao MSP. O MSP foi identificado utilizando o método proposto por (FERRARI; PINTO; MOREIRA, 2016). Esse espelhamento é um passo fundamental, pois garante que os extratores de características de forma e as representações aprendidas pela rede operem sobre estruturas anatômicas homólogas que compartilham uma orientação comum. Com isso, foram construídos dois vetores de características específicos para cada hemisfério, F_L para o hemisfério esquerdo e F_{R^*} para o hemisfério direito espelhado, compostos pelos seguintes atributos:

- ❑ **Atributos de Textura Aprendidos via DT-CWT (t):** Os mapas de energia da DT-CWT, derivados do hemisfério esquerdo e do hemisfério direito espelhado, serviram como entrada para os dois ramos da rede neural siamesa. A saída de cada ramo consiste em uma representação de características compacta e de alto nível, aprendida pela rede (t_L e t_{R^*}).
- ❑ **Descritor de Forma (*shape context*, h):** Um histograma log-esférico de 120 dimensões que codifica a geometria tridimensional do hipocampo. Este descritor foi calculado para a região hipocampal esquerda (h_L) e para a região hipocampal direita espelhada (h_{R^*}).
- ❑ **Atributos Volumétricos (v):** Um vetor tridimensional contendo os volumes escalares de GM, WM e CSF dentro do hipocampo. Esses valores foram calculados para o hipocampo esquerdo (v_L) e direito (v_R), normalizados pelo volume total de cada ROI. Como as medidas volumétricas são, por definição, invariantes à orientação espacial, o cálculo para o hemisfério direito não necessitou dos dados espelhados.

Dessa forma, os vetores de características compostos para o hemisfério esquerdo e direito espelhado foram formulados pela concatenação de seus respectivos atributos, conforme as Equações 66 e 67:

$$F_L = [t_L, h_L, v_L] \quad (66)$$

$$F_{R^*} = [t_{R^*}, h_{R^*}, v_R] \quad (67)$$

O vetor final de assimetria, $F_{\text{assimetria}}$, foi então derivado pelo cálculo da diferença absoluta, elemento a elemento, entre esses dois vetores compostos (Equação 68):

$$F_{\text{assimetria}} = \|F_L - F_{R^*}\| \quad (68)$$

Essa fusão resulta em um descritor unificado que captura a magnitude da diferença entre os hemisférios através de características aprendidas, de forma e volumétricas. O vetor de assimetria resultante serve como a entrada para as camadas de classificação finais da rede, permitindo que o modelo realize a predição diagnóstica com base em uma medida holística e multimodal da disparidade inter-hemisférica.

4.11 Aplicação do teste de McNemar

A metodologia proposta neste trabalho, que envolve a quantificação de assimetrias inter-hemisféricas para o treinamento de modelos de aprendizado de máquina em tarefas de classificação binária, requer uma análise estatística rigorosa para a comparação do desempenho dos classificadores. Dada a natureza pareada dos dados (o mesmo indivíduo é avaliado por diferentes abordagens de extração de atributos ou modelos) e o objetivo de comparar o desempenho em tarefas de classificação binária (e.g., CN vs. CCL), o teste de McNemar é uma ferramenta estatística apropriada para essa avaliação.

Conforme detalhado na Seção 2.12.1, o teste de McNemar é utilizado para avaliar a dependência de dados categóricos pareados. No contexto dos experimentos de classificação realizados, onde diferentes configurações de modelos são aplicadas ao mesmo conjunto de imagens para prever um resultado binário (e.g., CN ou CCL), as predições de cada modelo podem ser comparadas par a par. A consistência dos rótulos verdadeiros (*ground truth*) entre os modelos para cada tarefa e *fold* é validada, e as predições de todos os *folds* são agrupadas para cada tarefa.

Para cada par de modelos, uma tabela de contingência 2×2 é construída. Essa tabela sumariza as concordâncias e discordâncias entre as predições dos dois modelos em relação ao rótulo verdadeiro. As células b e c dessa tabela, que representam as discordâncias, são cruciais para o teste.

O teste de McNemar é então aplicado a cada par de modelos para calcular o valor de χ^2 e o p -valor correspondente, conforme descrito na Equação 47. Um nível de significância de $\alpha = 0.05$ é adotado. Um p -valor ≤ 0.05 indica que a diferença observada no desempenho dos modelos não é aleatória, mas sim resultado de uma diferença real entre eles. Essa lógica fornece uma interpretação estatística dos resultados, indicando a significância e qual modelo obteve melhor desempenho.

4.12 Considerações finais

Este capítulo apresentou a metodologia adotada para atingir os objetivos específicos indicados no Capítulo 1, Seção 1.2. Nesse sentido, este trabalho apresenta uma nova abordagem para identificar e quantificar assimetrias inter-hemisféricas em imagens de RM do cérebro, visando auxiliar o diagnóstico da DA. Vale ressaltar que essa abordagem é proposta como uma forma complementar para avaliação diagnóstica por imagem estrutural de RM, na medida que fornece informações relacionadas à assimetria inter-hemisférica que podem não ser facilmente identificadas, ou demandariam muito tempo de análise. No capítulo seguinte serão apresentados os resultados obtidos durante o desenvolvimento da pesquisa.

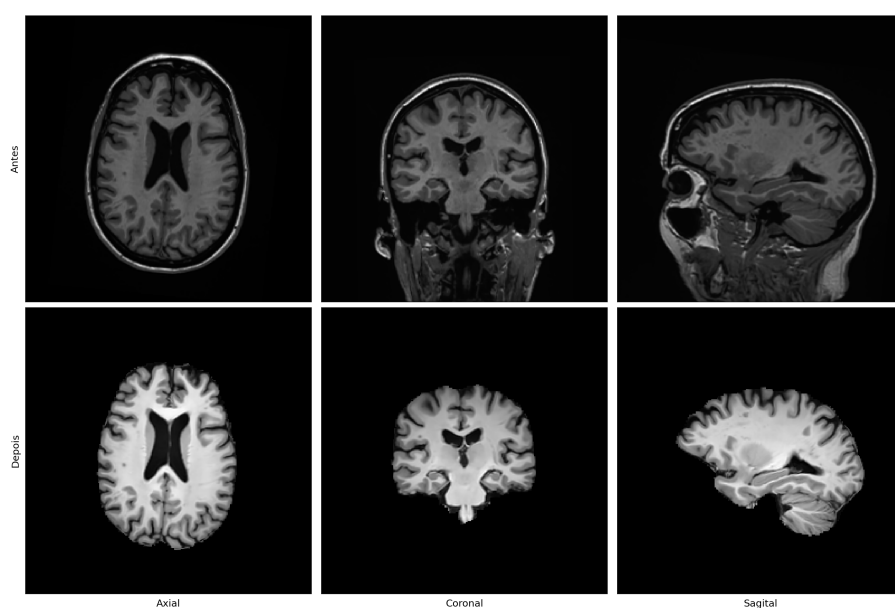
Capítulo 5

Resultados e discussões

5.1 Pré-processamento

Na Figura 5.1 é exibida uma comparação entre uma imagem original da base de dados ADNI antes e após as etapas de pré-processamento. As técnicas utilizadas foram descritas no Capítulo 4, sendo os resultados obtidos em cada etapa para essa imagem apresentados nos parágrafos a seguir.

Figura 19 – Visualização comparativa entre a imagem original (primeira linha) e a imagem preprocessada (segunda linha) nas vistas axial, coronal e sagital.

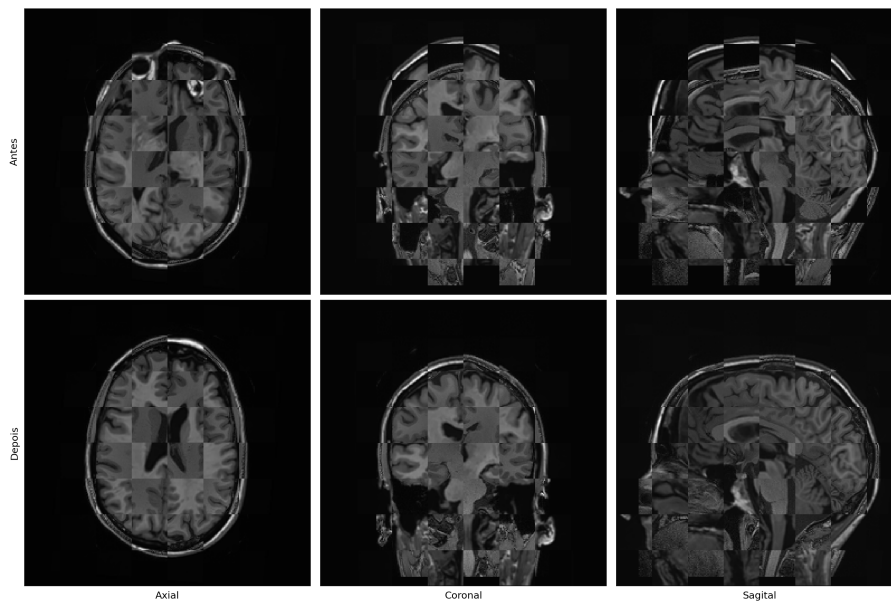


Fonte: elaborado pelo autor.

O correio afim tem como objetivo obter a transformação das imagens móveis (bases

IXI e ADNI) para que estejam alinhadas à imagem de referência SPL/NAC. O resultado apresentado na Figura 5.1 apresenta uma visualização no formato "tabuleiro de xadrez" que sobrepõe a imagem fixa e a imagem móvel de forma intercalada. O objetivo dessa visualização é avaliar visualmente a qualidade do corregristo, sendo um bom corregristo aquele com baixas variações entre as células do tabuleiro, de modo que a imagem resultante pareça ser contínua. Logo, nota-se que inicialmente estruturas anatômicas não apresentam nenhum alinhamento qualquer e, após o corregristo as principais estruturas do encéfalo apresentam certa continuidade. Embora o alinhamento não seja perfeito, esse não é um requisito da metodologia proposta uma vez que, o corregristo *b-spline* será utilizado para refinar esse alinhamento, deformando a imagem móvel para transferir os pontos estáveis para as imagens clínicas do estudo, sem que sejam alteradas as estruturas do encéfalo nas imagens analisadas.

Figura 20 – Resultado do corregristo afim entre a imagem de referência SPL/NAC e uma imagem da base ADNI.

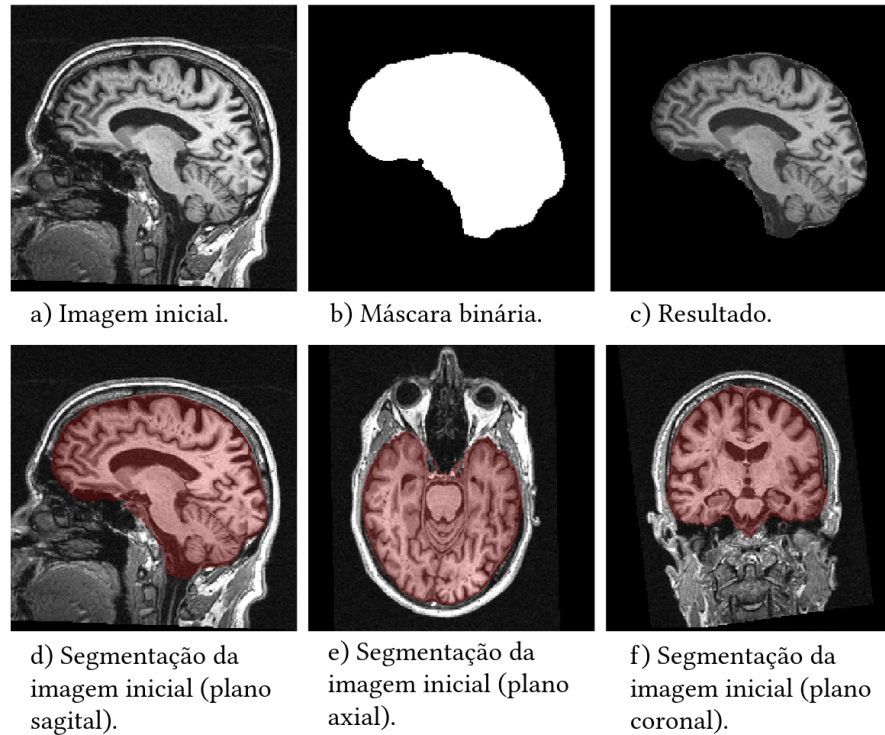


Fonte: elaborado pelo autor.

Após o alinhamento inicial, foi realizada a extração do encéfalo utilizando a técnica Synth-Strip. Na Figura 5.1, observou-se a remoção da caixa craniana e de outras estruturas não pertencentes ao cérebro após a aplicação da máscara binária do encéfalo na imagem de entrada.

Na sequência, realizou-se a redução de ruído com a técnica NLM. Na Figura 5.1, pode-se observar que o sinal original (primeira linha) apresentava variações de intensidade que não correspondiam a mudanças na estrutura anatômica, indicando que o sinal estava corrompido por ruído. Após a aplicação do filtro (segunda linha), essas variações foram drasticamente reduzidas. A parte corrompida do sinal foi visualizada por meio do cálculo da diferença absoluta entre a imagem original e a imagem filtrada (terceira linha).

Figura 21 – Resultado da extração do encéfalo com a técnica SynthStrip.



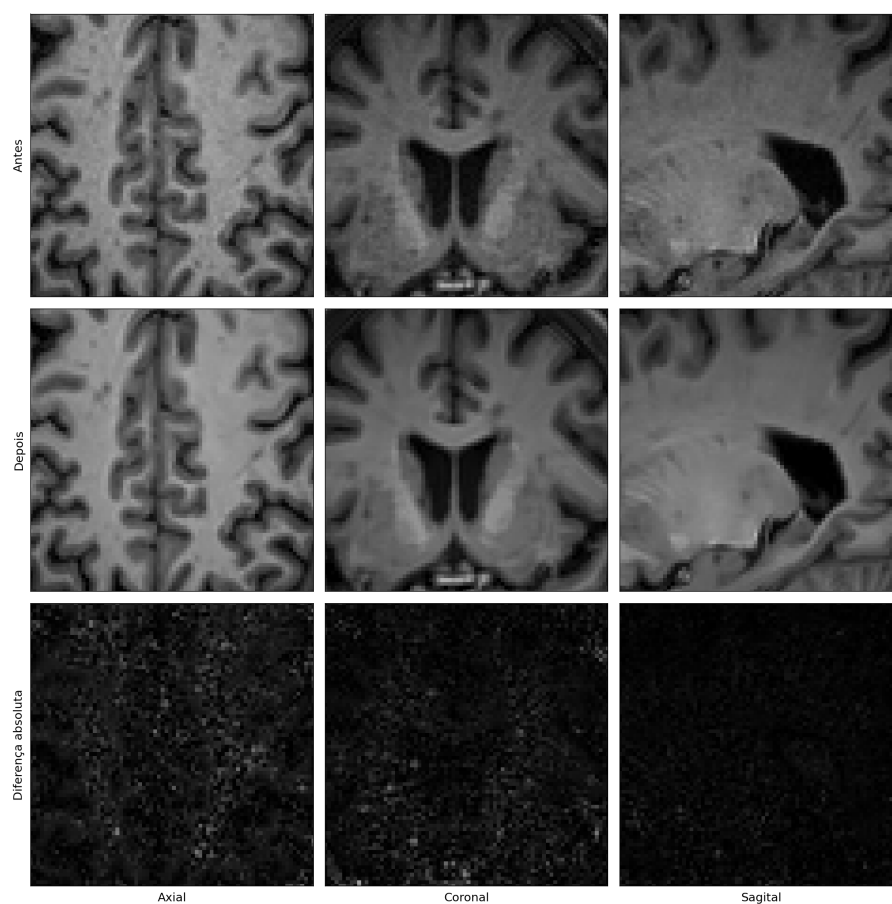
Fonte: elaborado pelo autor.

Após a remoção do ruído aditivo, foi realizada a correção do efeito *bias field*, que representava o componente multiplicativo no sinal original. A Figura 5.1 apresenta o resultado da correção com a técnica N4. Nessa figura, a diferença entre a imagem antes (a) e depois (b) da correção não foi visualizada tão facilmente. No entanto, ao analisar o *bias field* identificado (c), ficou clara a existência de variações não homogêneas de baixa frequência, características desse efeito. Na imagem (d), o *bias field* foi representado como uma superfície sob a imagem corrigida.

A remoção do ruído e do *bias field* eliminou os principais efeitos associados à degradação da imagem de RM. As etapas subsequentes tiveram como objetivo padronizar as imagens obtidas de diferentes indivíduos, *scanners*, hospitais e bases de dados, para reduzir as diferenças inerentes ao processo de aquisição. Com isso, esperava-se que as variações presentes nas imagens estivessem mais associadas às diferenças anatômicas entre indivíduos, ou à atrofia relacionada à DA, e menos a fatores como a potência do equipamento ou o posicionamento do indivíduo. Para isso, a imagem da base SPL/NAC foi utilizada como referência.

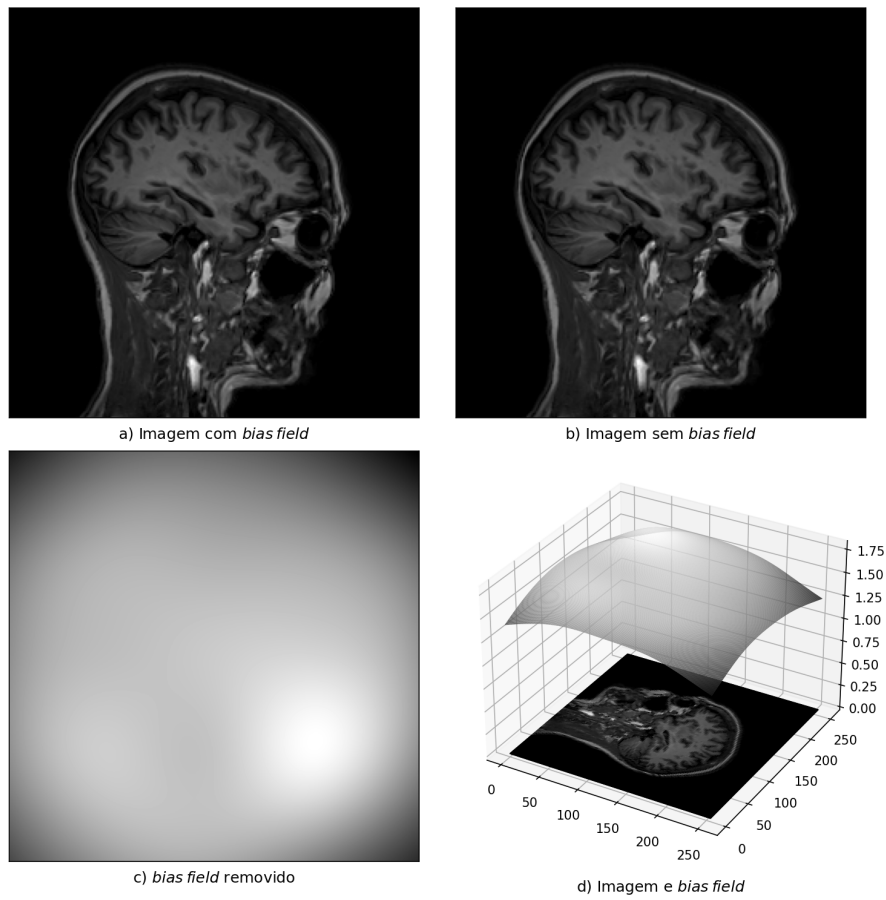
A última etapa consistiu na normalização de intensidades. Utilizando a técnica FCM, as imagens tiveram suas intensidades escalonadas com base na média da substância branca, conforme detalhado na Seção 2.5.3. A Figura 5.1 apresenta os histogramas de intensidade de diferentes imagens antes e após a normalização.

Figura 22 – Visualização comparativa para a redução de ruído com a técnica NLM em três vistas.



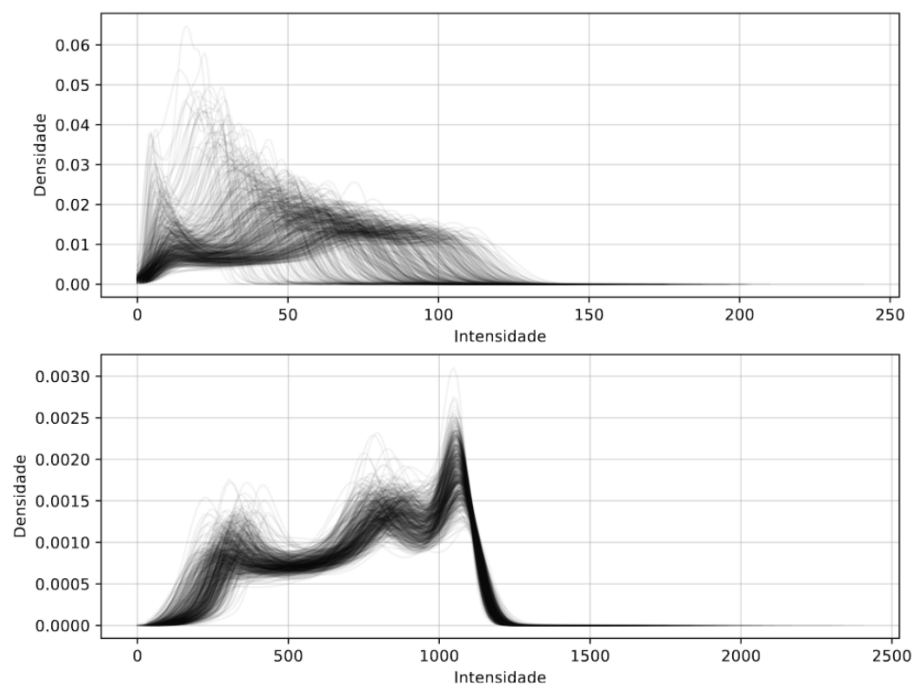
Fonte: elaborado pelo autor.

Figura 23 – Resultado da correção do efeito de *bias field* com a técnica N4.



Fonte: elaborado pelo autor.

Figura 24 – Exemplo comparativo de histogramas de intensidade antes (primeira linha) e depois (segunda linha) da normalização FCM, com $c_1 = 1000$.



Fonte: elaborado pelo autor.

5.2 Avaliação do detector de pontos salientes em imagens sintéticas 3D

Para uma análise inicial do detector de pontos salientes, foram conduzidos testes preliminares com imagens sintéticas 3D. Essas imagens, que representavam estruturas do tipo "ponta", foram geradas por meio do modelo paramétrico proposto por Wörz e Rohr (2006). Foram geradas 54 estruturas distintas, cada uma com 12 versões, sendo que cada versão apresentava um nível de degradação causado por ruído e *bias field*, totalizando 648 imagens. As imagens possuíam dimensões de $64 \times 64 \times 64$ voxels, e a localização do ponto saliente de referência foi calculada pelo modelo para que estivesse sempre no centro de cada imagem. Assim, os pontos detectados fora da região de interesse ou os pontos adicionais dentro dessa região foram considerados falsos positivos (FP), enquanto os falsos negativos (FN) referiram-se aos casos em que nenhum ponto foi detectado na região de interesse.

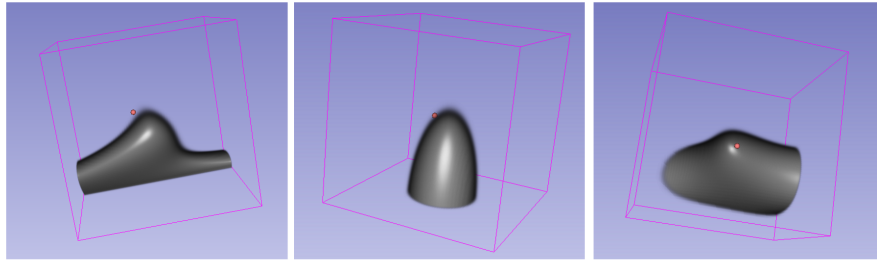
Uma busca em grade foi utilizada para determinar os valores ideais dos parâmetros α , β e T_E , conforme descrito na Sub-seção 4.5. Os seguintes intervalos de valores foram considerados: $\alpha \in \{0.25, 0.5, 0.75, 1\}$, $\beta \in \{1/28, 0.05, 0.1\}$ e $T_E \in \{0.5, 0.75\}$. A região de interesse foi definida como uma esfera com origem no voxel central da imagem e raio de 10 voxels. O número de escalas e o tamanho da vizinhança foram mantidos fixos em 3 e $5 \times 5 \times 5$, respectivamente. Os melhores resultados foram obtidos com $\alpha = 1$, $\beta = 0,05$ e $T_E = 0,75$, valores consistentes com os parâmetros utilizados por Souza, Poloni e Ferrari (2020). Apenas as imagens sem adição de ruído ou *bias field* foram utilizadas nessa busca. A Figura 5.2 ilustra algumas das estruturas sintéticas utilizadas e os pontos detectados (em vermelho) pela técnica DT-CWT. Esses parâmetros foram mantidos, e uma nova detecção foi realizada adicionando as imagens que sofreram degradações. O objetivo desta etapa foi avaliar a robustez do detector a essas variações. Os resultados são apresentados na Tabela 5.2 e indicaram que o detector DT-CWT 3D foi eficaz na identificação de pontos de alta curvatura. Contudo, a aplicação de técnicas de remoção de ruído e correção de *bias field* mostrou-se essencial para reduzir a quantidade de falsos positivos.

Tabela 9 – Valores médios da distância (d) entre o ponto ideal e o ponto detectado e quantidade de FPs e FNs que ocorreram nos conjuntos, avaliados em crescentes níveis de degradação com ruído (σ_r) e *bias field* (BF).

	$BF = 0\%$			$BF = 10\%$		
	d	FP	FN	d	FP	FN
$\sigma_r = 0\%$	$3,81 \pm 0,74$	2	0	$5,51 \pm 0,79$	3	2
$\sigma_r = 1\%$	$3,83 \pm 0,75$	2	0	$5,48 \pm 0,81$	3	1
$\sigma_r = 2\%$	$4,57 \pm 0,81$	2	0	$6,80 \pm 0,88$	4	19
$\sigma_r = 3\%$	$4,61 \pm 0,75$	4	1	$6,89 \pm 0,88$	6	15

Fonte: elaborada pelo autor.

Figura 25 – Pontos detectados em três imagens sintéticas geradas por um modelo paramétrico.



Fonte: elaborado pelo autor.

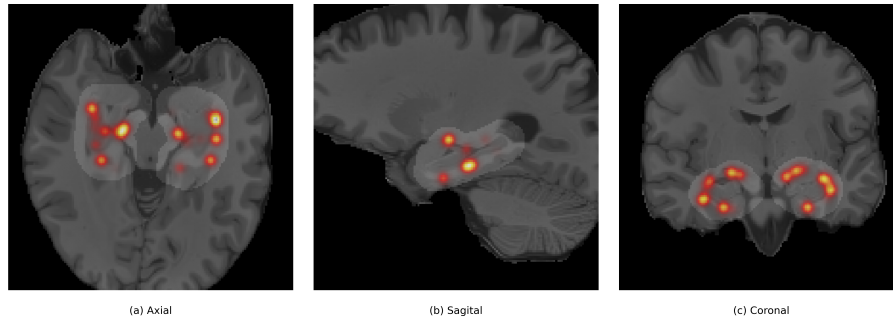
5.3 Detecção de pontos estáveis na população CN

O objetivo deste experimento foi analisar as imagens de indivíduos na faixa etária de 60 a 69 anos, que correspondia ao início do diagnóstico de CCL nas imagens selecionadas da base ADNI. Para a execução, foram utilizadas 118 imagens da base IXI de indivíduos com idades na faixa etária de $[60, 70)$ anos. Após o corregristo com a imagem de referência SPL/NAC, realizou-se a detecção dos pontos salientes no hemisfério esquerdo e no hemisfério direito rotacionado, utilizando as máscaras da região hipocampal para restringir a detecção a essa área. Os parâmetros $\alpha = 1$, $\beta = 0,05$ e $T_E = 0,75$, obtidos nos experimentos anteriores, e os valores de três escalas e vizinhança de tamanho $5 \times 5 \times 5$ foram utilizados.

Na sequência, a técnica MPW foi usada para construir uma nova imagem volumétrica representando a distribuição espacial dos pontos salientes em cada hemisfério, com a intensidade de cada *voxel* correspondendo à densidade de pontos em sua localização. A Figura 5.3 mostra três projeções do atlas probabilístico de pontos salientes 3D. Ao analisar visualmente essas imagens, observou-se que determinadas regiões apresentavam uma densidade significativamente maior em relação à sua vizinhança, indicando que essas posições tinham o potencial de representar os pontos salientes estáveis nesta população. A etapa subsequente calculou as posições dos máximos locais, considerando uma vizinhança de tamanho $5 \times 5 \times 5$ *voxels*. O resultado foi um conjunto de pontos salientes estáveis, com

25 pontos no hemisfério direito (rotacionado) e 29 no hemisfério esquerdo. Para permitir uma comparação direta entre os hemisférios, os 4 pontos do hemisfério esquerdo com menor intensidade foram removidos, resultando em 25 pontos em ambos os hemisférios.

Figura 26 – Projeções axial, coronal e sagital do atlas probabilístico criado para as regiões hipocampais, a partir dos pontos salientes detectados na população CN.



Fonte: elaborado pelo autor.

5.4 Posicionamento do Atlas nas Imagens da Base ADNI

Conforme descrito na Seção 2.8, o mapa de rótulos de segmentação das regiões de interesse foi alinhado às imagens clínicas via transformação *B-spline* deformável, sendo a imagem da base ADNI definida como a parte fixa da transformação. Logo, foi possível reaplicar a mesma transformação, salva em arquivo, para posicionar o atlas sobre as imagens clínicas, não sendo necessário realizar um novo alinhamento. Uma vez alinhados espacialmente, em uma segunda etapa, os pontos foram correlacionados para encontrar as coordenadas de máximo local no mapa de energia acumulada, utilizando uma janela de $10 \times 10 \times 10$ *voxels*, centralizada em cada ponto transformado, e um *kernel* de $3 \times 3 \times 3$. Juntas, essas duas etapas foram empregadas visando garantir o melhor posicionamento dinâmico dos pontos, considerando possíveis pequenos erros no corregristo e adaptando-se à região de alta curvatura local e relativa a cada indivíduo. O refinamento e o posicionamento foram realizados sempre considerando a máscara da estrutura anatômica em questão, evitando que pontos fossem definidos em uma estrutura vizinha.

5.5 Definição dos Parâmetros do *Shape Context*

O propósito deste experimento foi determinar qual das quatro estratégias de posicionamento do descritor *shape context* gera a métrica de assimetria inter-hemisférica com o maior poder de discriminação entre as populações CN e DA. A avaliação foi conduzida utilizando o conjunto de dados de otimização, e a estratégia ótima foi definida como

aquela que maximiza a separabilidade estatística entre as distribuições de assimetria dos dois grupos.

Uma vez obtidos os descritores, foi calculada a distância χ^2 , Equação 38, entre cada hemisfério. Os valores utilizados na busca em grade foram: número de escalas para o mapa de energia acumulada $m = \{1, 2, 3\}$; ponderador de escala $\alpha = \{0.25, 0.5, 1.0\}$; ponderador de orientação $\beta = \{1/28, 0.25, 0.5, 1\}$; número de divisões radiais $n_r = \{3, 4, 5, 6\}$; número de divisões de escala $n_e = \{3, 4, 6, 8\}$; número de divisões de azimuth $n_a = \{4, 6, 8, 12\}$; e peso da função Gaussiana $\sigma = \{0.5, 0.75, 1.0, 1.5\}$.

Para quantificar a separabilidade, foi utilizado o Índice de Fisher (IZENMAN, 2013), cuja fórmula é dada por:

$$F = \frac{(\mu_{DA} - \mu_{CN})^2}{\sigma_{DA}^2 + \sigma_{CN}^2} \quad (69)$$

onde μ e σ^2 representam, respectivamente, a média e a variância da métrica de assimetria para cada grupo diagnóstico. Um valor mais alto de F indica maior separabilidade.

A Tabela 10 sumariza os resultados quantitativos, enquanto a Figura 27 ilustra as distribuições de densidade para cada estratégia.

Tabela 10 – Análise de Separação de Classe para as Estratégias de Posicionamento do Descritor.

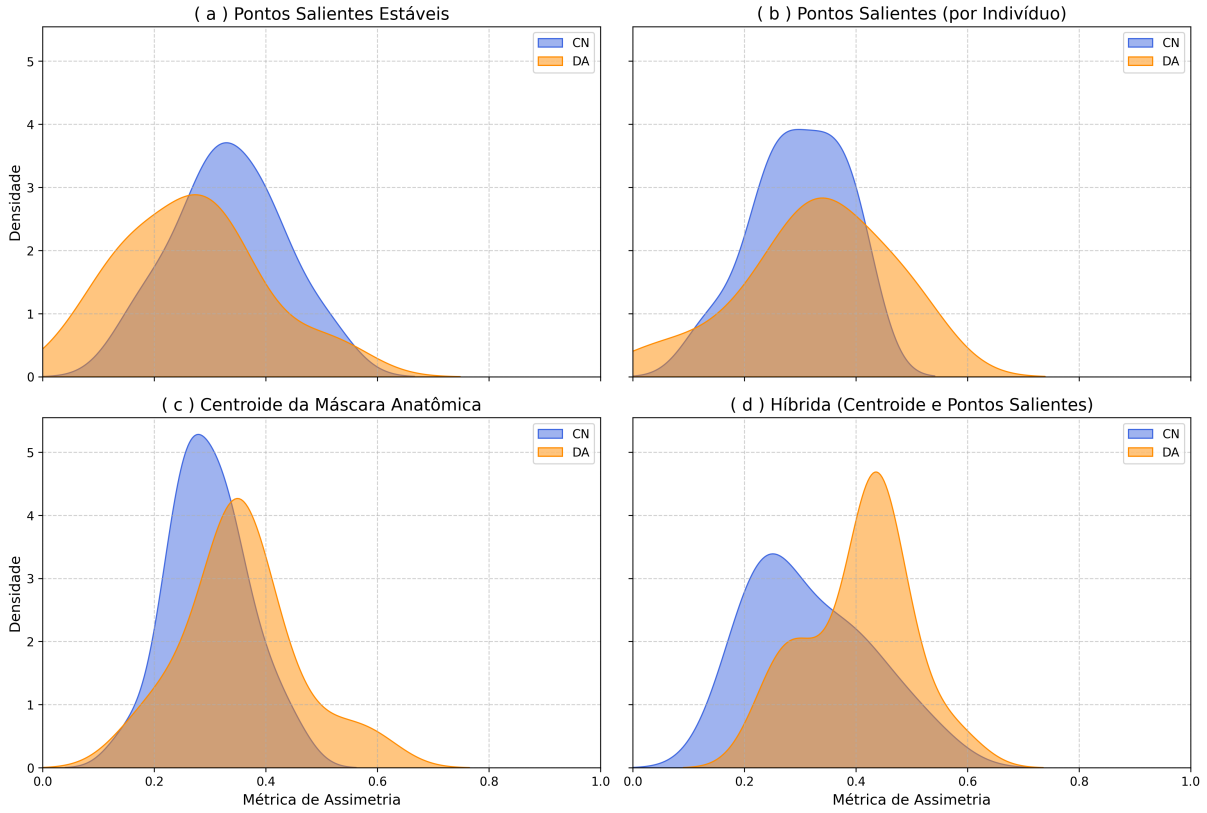
Estratégia de Posicionamento	Índice de Separação (Fisher)	Valor-p (Teste t)
1. Pontos Salientes Estáveis	0,1911	0,0643
2. Pontos Salientes (por Indivíduo)	0,0383	0,3987
3. Centroide da Máscara Anatômica	0,2230	0,0464
4. Híbrida (Centroide e Pontos Salientes)	0,3977	0,0091

Fonte: elaborado pelo autor.

A Tabela 10 demonstra que a estratégia baseada em **Centroide e Pontos Salientes** alcançou o maior Índice de Separação de Fisher (0,3977), indicando a máxima separabilidade entre os grupos CN e DA. Este resultado quantitativo é visualmente corroborado pela Figura 27(d), que exibe a separação entre as distribuições de assimetria das duas populações, com mínima sobreposição em comparação com as demais estratégias. Adicionalmente, esta abordagem resultou no menor p -valor (0,0091) no teste- t , confirmando a alta significância estatística da diferença entre os grupos.

Os melhores resultados foram obtidos para a combinação: $m = 2$, $\alpha = 1$, $n_r = 5$, $n_e = 4$, $n_a = 6$ e $\sigma = 0,75$. A busca foi então refinada para o parâmetro $\beta = \{1/28, 0.05, 0.1, 0.15, 0.2, 0.25\}$, que resultou no valor final de $\beta = 0,05$. Portanto, cada histograma *shape context* passou a ser definido por um vetor de 120 atributos.

Figura 27 – Distribuições de densidade da métrica de assimetria inter-hemisférica para os grupos CN e DA, para cada uma das quatro estratégias de posicionamento do descritor.



Fonte: elaborado pelo autor.

5.6 Resultados da classificação

A metodologia proposta foi avaliada em quatro tarefas de classificação binária: CN vs. CCL, CN vs. DA, CCL vs. DA e CCLe vs. CCLp. Cada tarefa investiga diferentes estágios do contínuo da doença de Alzheimer, desde o declínio cognitivo inicial até a neurodegeneração confirmada. Para todos os experimentos, os resultados são reportados em termos de sensibilidade (SEN), especificidade (SPE), acurácia balanceada (BACC) e F1-score (SOKOLOVA; LAPALME, 2009). O impacto de cada conjunto de características (*shape context*, volume tecidual e textura baseada em wavelet) é analisado através de estudos pontuais com a adição de cada técnica.

5.6.1 Configuração Experimental

Estratégia de Validação Cruzada. Para garantir uma avaliação robusta e imparcial, todos os experimentos foram realizados usando um esquema de validação cruzada de 5 *folds*. Os *folds* foram estratificados para manter uma distribuição de classes balanceada em cada divisão. Além disso, para evitar o vazamento de dados das amostras longitu-

dinais do conjunto de dados ADNI (onde existem múltiplas varreduras para um único sujeito), foi aplicada uma estratégia de agrupamento por sujeito. Isso garantiu que todas as varreduras de um determinado sujeito fossem confinadas a um único fold, seja para treinamento ou para teste, mas nunca divididas entre eles.

Aumento de Dados. Para aumentar a diversidade dos dados de treinamento e melhorar a generalização do modelo, foi empregado um extenso pipeline de aumento de dados usando o *framework* MONAI. Para cada *epoch*, 50% das amostras de treinamento foram selecionadas aleatoriamente para passar por uma série de transformações, aplicadas com a mesma intensidade em cada hemisfério mas com *seeds* distintas. Nesse sentido, embora a remoção de ruído já tenha sido realizada, essa etapa incluiu a adição de novos tipos de ruídos, que são inseridos de maneira deliberada e controlada. Cada técnica de aumento foi aplicada com uma probabilidade variando de 10% a 50%, garantindo que cada amostra selecionada recebesse pelo menos uma transformação. Estas incluíam:

- ❑ **Aumentos Espaciais:** Rotação aleatória, zoom, transformação afim e deformação elástica.
- ❑ **Aumentos de Intensidade:** Aplicação de ruído Gaussiano, suavização Gaussiana, ajustes de contraste, deslocamentos e escalonamentos de intensidade, campos de bias, ruído de Gibbs e ruído de pico no espaço K.

Os modelos de aprendizado profundo foram treinados com 50 épocas, com parada antecipada definida para uma perda mínima e delta de 0,001, e paciência de 5 épocas. A taxa de aprendizado foi de 0,001 com decaimento de peso de 0,0001 e otimizador Adam, e 8 amostras por *batch*. Os parâmetros do classificador XGBoost foram definidos como `n_estimators = 200`, `max_depth = 5`, `learning_rate = 0,1`, `colsample_bytree = 0,9` e `subsample = 0,7`.

5.7 Classificação CN vs. CCL

A distinção entre o envelhecimento saudável e os estágios iniciais da neurodegeneração, representados pelo CCL, é fundamental para o diagnóstico precoce. Contudo, esta tarefa é notoriamente difícil devido às alterações cerebrais sutis e heterogêneas que caracterizam a fase inicial da DA.

Os resultados quantitativos para este cenário estão sumarizados na Tabela 11. O modelo de referência (Baseline), que utiliza apenas as imagens de RM da região hipocampal como entrada para uma rede CNN Siamesa, alcançou um desempenho modesto, com um F1-score de 0,662 e uma acurácia balanceada de 0,651. Este resultado serve como ponto de partida para avaliar o impacto da incorporação de informações de assimetria.

Uma melhora substancial foi observada ao integrar a SCNN com o mapa de energia (t) e descritores Shape Context (h). Esta configuração elevou o F1-score para **0,757** e

a BACC para **0,760**, um ganho significativo em relação a todas as outras abordagens. Tal avanço valida a hipótese central deste trabalho: as discrepâncias geométricas inter-hemisféricas, capturadas pelos descritores de forma e potencializadas pela arquitetura Siamesa, funcionam como biomarcadores eficazes para detectar as primeiras manifestações patológicas, mesmo antes que a atrofia volumétrica se torne pronunciada.

As curvas ROC, apresentadas na Figura 28, ilustram visualmente a superioridade discriminativa do modelo SCNN. É interessante notar que a adição de características de volume tecidual ao modelo (h e v) não resultou em uma melhoria adicional, levando a uma leve queda no desempenho (F1-score de 0,734). Este resultado sugere que, nesta fase inicial da doença, as informações de volume podem ser redundantes ou até mesmo introduzir ruído, uma vez que as alterações de forma já capturam a essência da degeneração incipiente. Nesse sentido, as métricas de volume total, i.e., ambos os hemisférios juntos, podem ser consideradas mais representativas na análise do grupo CCL, como foi observado por Ocklenburg et al. (2024), Matsuda et al. (2025).

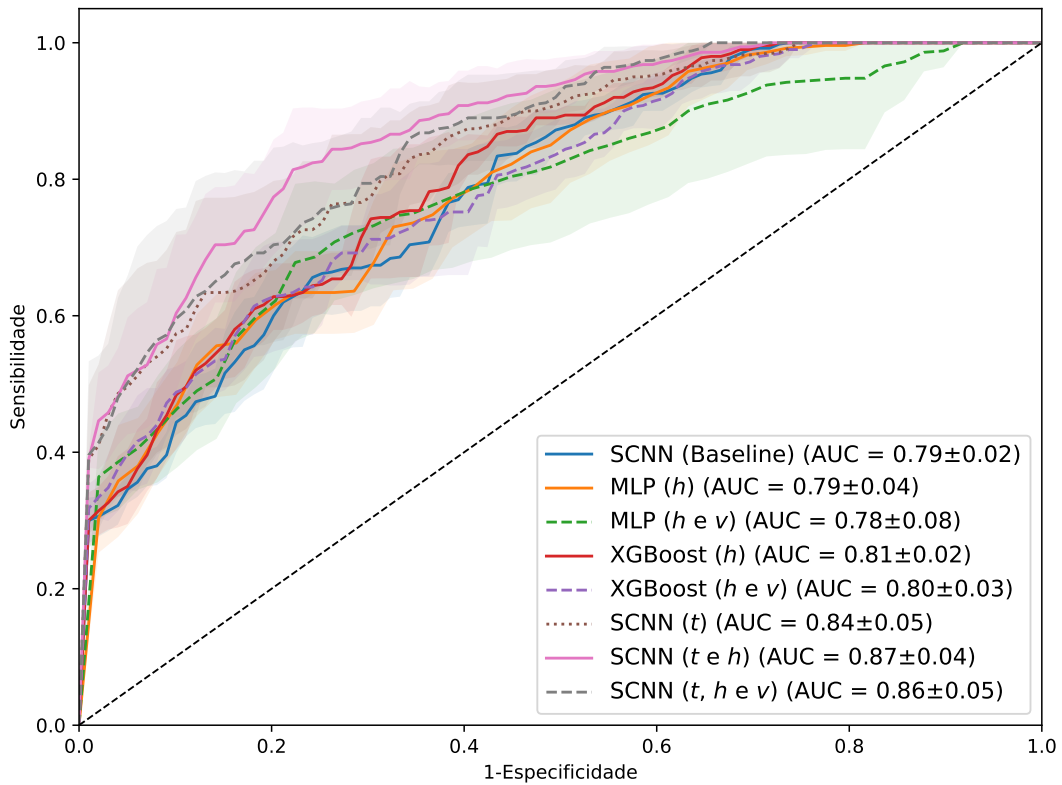
O desempenho obtido mostra-se competitivo em relação aos trabalhos correlatos. Por exemplo, Poloni et al. (2021), utilizando atributos inter-hemisféricos de textura na região hipocampal, reportaram uma acurácia de 69,44% e AUC de 0,76, enquanto o uso de métricas volumétricas resultou em 72,68% de acurácia e 0,8 de AUC. O resultado superior obtido no presente estudo sugere que os modelos multi-modais são mais eficazes em capturar as alterações estruturais na região hipocampal nesses grupos diagnósticos. Nesse sentido, esses resultados corroboram as observações de Matsuda et al. (2025), que constataram que, o volume total do hipocampo e amígdala é mais relevante que a assimetria volumétrica em indivíduos com comprometimento cognitivo. Isso reforça a hipótese de que, na fase inicial, as deformações geométricas e as alterações texturais precedem uma atrofia volumétrica significativa, tornando-se biomarcadores mais sensíveis.

Tabela 11 – Resultados da classificação entre indivíduos CN e com CCL.

Modelo	Atributos	SEN	SPE	BACC	F1-score
SCNN (Baseline)	Imagem	$0,686 \pm 0,06$	$0,641 \pm 0,01$	$0,651 \pm 0,02$	$0,662 \pm 0,03$
XGBoost	h	$0,654 \pm 0,04$	$0,669 \pm 0,04$	$0,665 \pm 0,04$	$0,661 \pm 0,03$
MLP	h	$0,616 \pm 0,05$	$0,686 \pm 0,04$	$0,667 \pm 0,04$	$0,648 \pm 0,04$
XGBoost	h e v	$0,662 \pm 0,06$	$0,656 \pm 0,04$	$0,656 \pm 0,04$	$0,657 \pm 0,04$
MLP	h e v	$0,678 \pm 0,03$	$0,674 \pm 0,06$	$0,672 \pm 0,04$	$0,675 \pm 0,03$
SCNN	t	$0,720 \pm 0,05$	$0,724 \pm 0,08$	$0,720 \pm 0,06$	$0,721 \pm 0,06$
SCNN	t e h	$0,746 \pm 0,02$	$0,771 \pm 0,05$	$0,760 \pm 0,03$	$0,757 \pm 0,03$
SCNN	t , h e v	$0,728 \pm 0,07$	$0,742 \pm 0,08$	$0,736 \pm 0,07$	$0,734 \pm 0,07$

Fonte: elaborado pelo autor.

Figura 28 – Curvas ROC e seus respectivos valores de AUC para a tarefa de classificação CN vs. CCL.



Fonte: elaborado pelo autor.

5.8 Classificação CN vs. DA

A classificação entre CN e DA é um problema mais tratável, pois a DA é caracterizada por uma atrofia cerebral acentuada e bem documentada, especialmente no lobo temporal medial. Como esperado, todos os modelos propostos superaram o desempenho obtido no cenário CN vs. CCL.

Conforme detalhado na Tabela 12, o modelo SCNN completo (t , h e v) alcançou o melhor desempenho geral, com um F1-score de 0,872 e uma BACC de 0,866. Diferentemente do cenário anterior, aqui a adição de características de volume tecidual foi benéfica, sugerindo que em estágios mais avançados da doença, a perda de volume e as alterações de forma são fontes de informação complementares. Enquanto a assimetria de forma captura as deformações geométricas, o volume quantifica a magnitude da atrofia regional, e a textura (via DT-CWT) detecta a degeneração do tecido.

A inclusão de atributos multi-modais melhorou significativamente o desempenho do modelo SCNN. Em sua versão com atributos de textura, o modelo apresentou um desempenho robusto (F1-score de 0,858). A adição dos descritores de forma e volume melhorou ainda mais a capacidade do modelo, resultando em um classificador com a maior sensibilidade (0,908). Nesse sentido, os atributos multi-modais são cruciais na identificação

correta de pacientes com DA. A Figura 5.8 apresenta o gráfico ROC e valores AUC para os modelos avaliados.

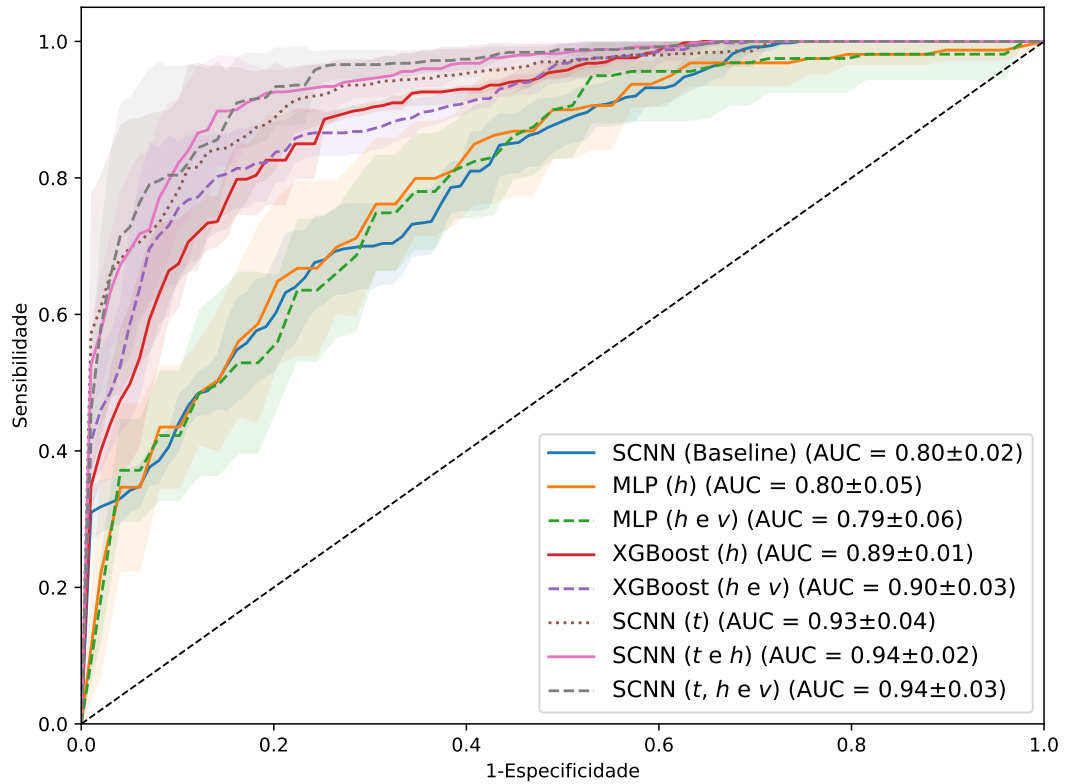
Estes resultados estão alinhados com os estudos de Oliveira, Poloni e Ferrari (2020) e Poloni et al. (2021), que apontam a assimetria hipocampal como um forte indicador de DA, alcançando acurácia de 81,43% e 82,59%, respectivamente. A contribuição principal deste estudo reside em demonstrar que a fusão de múltiplas características estruturais (forma, volume e textura) dentro de uma rede Siamesa para análises de assimetrias inter-hemisféricas na região hipocampal é capaz de identificar os diferentes estágios da DA. Enquanto os trabalhos anteriores validaram a relevância da assimetria hipocampal, este estudo demonstra que a classificação multi-modal melhora a capacidade discriminativa do modelo. O desempenho é também competitivo com abordagens de aprendizado profundo mais recentes, como a de Li, Ji e Sun (2024), que alcançou um F1-score de 0,903 utilizando uma arquitetura de aprendizado profundo com pontos salientes de todo o cérebro.

Tabela 12 – Resultados da classificação entre indivíduos CN e com DA.

Modelo	Atributos	SEN	SPE	BACC	F1-score
SCNN (Baseline)	Imagem	0,716 $\pm$ 0,06	0,651 $\pm$ 0,01	0,666 $\pm$ 0,02	0,681 $\pm$ 0,03
XGBoost	<i>h</i>	0,754 $\pm$ 0,02	0,796 $\pm$ 0,03	0,780 $\pm$ 0,02	0,774 $\pm$ 0,02
MLP	<i>h</i>	0,649 $\pm$ 0,2	0,770 $\pm$ 0,07	0,707 $\pm$ 0,06	0,696 $\pm$ 0,08
XGBoost	<i>h</i> e <i>v</i>	0,772 $\pm$ 0,06	0,851 $\pm$ 0,06	0,816 $\pm$ 0,04	0,807 $\pm$ 0,04
MLP	<i>h</i> e <i>v</i>	0,756 $\pm$ 0,01	0,730 $\pm$ 0,05	0,713 $\pm$ 0,05	0,738 $\pm$ 0,05
SCNN	<i>t</i>	0,870 $\pm$ 0,06	0,848 $\pm$ 0,08	0,855 $\pm$ 0,06	0,858 $\pm$ 0,06
SCNN	<i>t</i> e <i>h</i>	0,876 $\pm$ 0,02	0,860 $\pm$ 0,05	0,865 $\pm$ 0,03	0,867 $\pm$ 0,03
SCNN	<i>t</i> , <i>h</i> e <i>v</i>	0,908 $\pm$ 0,05	0,840 $\pm$ 0,06	0,866 $\pm$ 0,05	0,872 $\pm$ 0,05

Fonte: elaborado pelo autor.

Figura 29 – Curvas ROC para a classificação CN vs. DA.



Fonte: elaborado pelo autor.

5.9 Classificação CCL vs. DA

A distinção entre CCL e DA é desafiadora porque representa um contínuo de neurodegeneração, com sobreposição significativa nos padrões de atrofia. O objetivo aqui é identificar o ponto de transição em que o comprometimento cognitivo se consolida como demência.

Os resultados, apresentados na Tabela 13, mostram que, embora o modelo SCNN com todos atributos (t , h e v) tenha novamente obtido o melhor desempenho (F1-score de 0,731), os ganhos em relação aos modelos mais simples foram mais modestos. A rede Siamesa sozinha já alcança um F1-score de 0,706, e a adição de descritores de forma e volume confere melhorias incrementais.

Isso sugere que, ao comparar estágios adjacentes da doença, o poder discriminativo adicional oferecido pelas características de forma e volume é limitado. Muitos indivíduos com CCL já apresentam atrofia considerável, tornando a fronteira de decisão entre as classes menos definida. As distribuições de características se sobrepõem, dificultando a classificação. Como visto na Figura 30, as curvas ROC estão mais próximas umas das outras em comparação com o cenário CN vs. DA, refletindo a maior dificuldade da tarefa.

Embora o modelo completo tenha alcançado a maior sensibilidade (0,758), a melhoria na especificidade foi contida. Este resultado aponta para a heterogeneidade da população

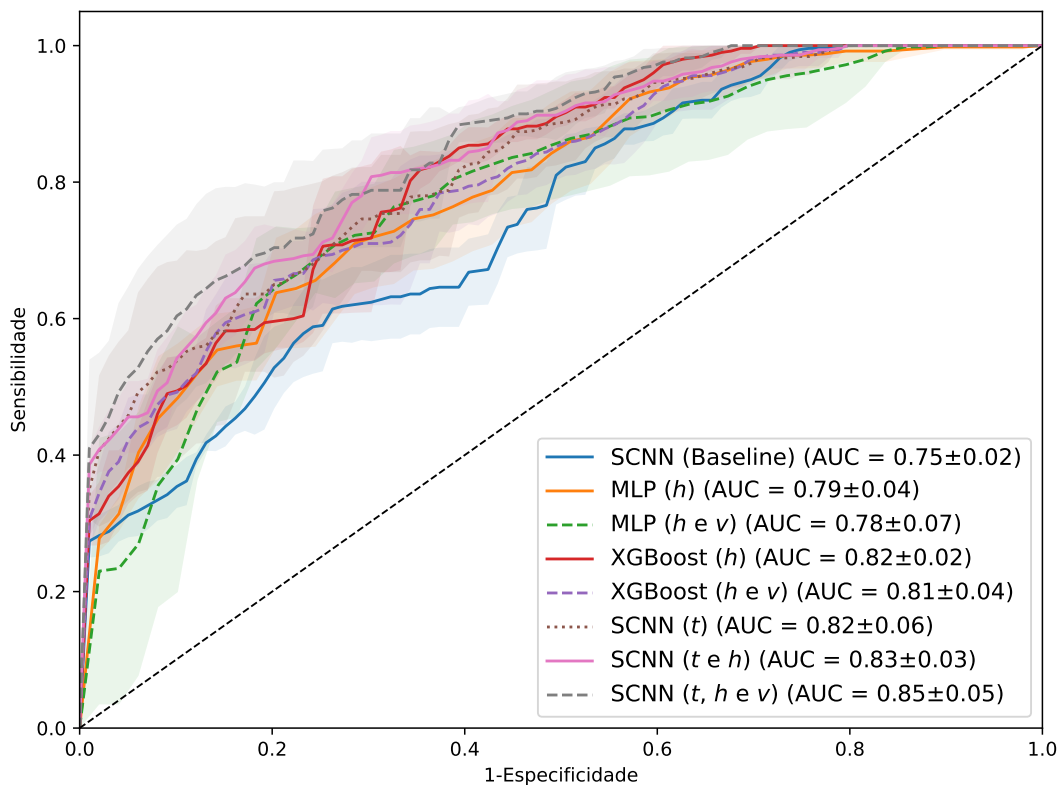
com CCL e sugere a necessidade de explorar biomarcadores mais dinâmicos, como dados longitudinais (análise de imagens ao longo do tempo) ou de conectividade funcional, para capturar a transição de CCL para DA de forma mais robusta.

Tabela 13 – Resultados da classificação CCL vs. DA.

Modelo	Atributos	SEN	SPE	BACC	F1-score
SCNN (Baseline)	Imagem	$0,646 \pm 0,06$	$0,592 \pm 0,02$	$0,601 \pm 0,02$	$0,617 \pm 0,03$
XGBoost	h	$0,604 \pm 0,04$	$0,687 \pm 0,04$	$0,664 \pm 0,03$	$0,642 \pm 0,04$
MLP	h	$0,649 \pm 0,09$	$0,765 \pm 0,06$	$0,707 \pm 0,06$	$0,696 \pm 0,08$
XGBoost	h e v	$0,642 \pm 0,06$	$0,685 \pm 0,05$	$0,671 \pm 0,04$	$0,661 \pm 0,04$
MLP	h e v	$0,698 \pm 0,05$	$0,700 \pm 0,05$	$0,697 \pm 0,04$	$0,698 \pm 0,03$
SCNN	t	$0,740 \pm 0,05$	$0,676 \pm 0,07$	$0,690 \pm 0,06$	$0,706 \pm 0,06$
SCNN	t e h	$0,706 \pm 0,02$	$0,700 \pm 0,04$	$0,700 \pm 0,03$	$0,702 \pm 0,02$
SCNN	t , h e v	$0,758 \pm 0,06$	$0,707 \pm 0,05$	$0,721 \pm 0,05$	$0,731 \pm 0,05$

Fonte: elaborado pelo autor.

Figura 30 – Curvas ROC para a classificação CCL vs. DA.



Fonte: elaborada pelo autor.

5.10 Classificação CCLe vs. CCLp

Identificar quais indivíduos com CCL irão progredir para DA (CCLp) em um futuro próximo é, talvez, o maior desafio e a tarefa de maior impacto clínico na pesquisa de

neuroimagem da DA. Esta classificação prognóstica depende da detecção de alterações extremamente sutis que predizem a conversão.

Os resultados, apresentados na Tabela 14, demonstram a relevância da abordagem multimodal neste cenário de classificação. O modelo SCNN completo (t , h e v) alcançou um bom desempenho, com uma BACC de 0,841 e um F1-score de 0,844. Somado a isso, este modelo apresentou um alto equilíbrio entre sensibilidade (0,858) e especificidade (0,832), o que é ideal para aplicações clínicas, pois minimiza tanto os falsos negativos (não identificar um paciente que irá progredir) quanto os falsos positivos (alarmar desnecessariamente um paciente estável).

Nesta tarefa de classificação, a análise da contribuição de cada componente revela que a assimetria geométrica (h) aplicada no classificador XGBoost não foi capaz de diferenciar os dois grupos de maneira eficiente (F1-score de 0,650). No entanto, a arquitetura SCNN, ao aprender padrões de assimetria com atributos de textura, elevou significativamente o desempenho (F1-score de 0,805). A combinação SCNN (t e h) obteve o melhor resultado, com F1-score de 0,845.

Isso sugere que a progressão da doença é marcada por uma combinação de fatores na região hipocampal: alterações sutis na textura dos tecidos (extraídas pela DT-CWT), deformações geométricas assimétricas (capturadas pelo *shape context*) e o início da atrofia volumétrica (capturada pelas métricas de segmentação tecidual). A fusão dessas três fontes de informação é, portanto, essencial para um prognóstico preciso. A Figura 31 apresenta os valores de AUC, mostrando um claro ganho de desempenho com a integração das modalidades.

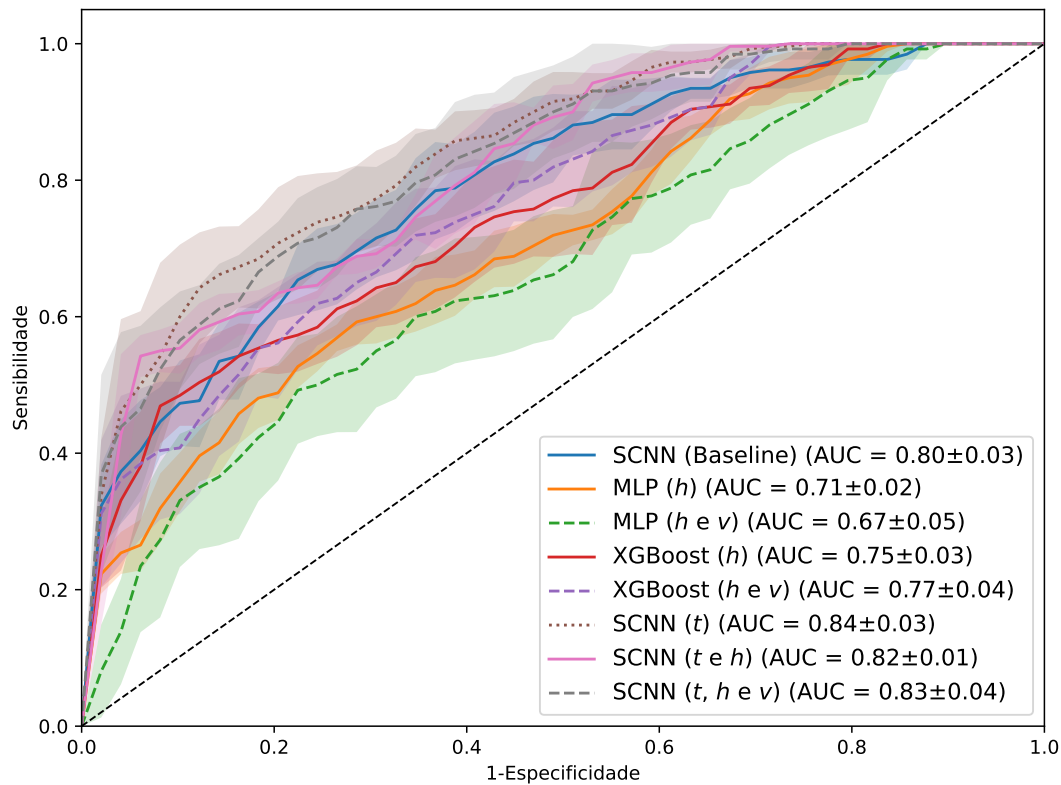
O desempenho obtido é superior aos modelos de aprendizado profundo baseados em RM, como o de Luo et al. (2023) (F1-score de 0,75) e o de Liu et al. (2025) (F1-score de 0,796). Além disso, é comparável com o modelo de Luo et al. (2024) (F1-score de 0,82) e Agostinho, Simões e Castelo-Branco (2024) (F1-score de 0,84). O sucesso da abordagem alinha-se conceitualmente com os estudos pioneiros de Long, Jiang e Zhang (2018) e Wachinger et al. (2016), que já apontavam a assimetria morfológica e de forma em regiões límbicas como um biomarcador chave para a progressão da doença. Entretanto, os estudos de Rahim et al. (2025) (91,77% e AUC de 92,87%) e Niyas e Thiyagarajan (2023) (acurácia de 96%) demonstram que ainda existem lacunas no entendimento do contínuo da DA para serem exploradas.

Tabela 14 – Resultados da classificação prognóstica entre CCLe e CCLp.

Modelo	Atributos	SEN	SPE	BACC	F1-score
SCNN (Baseline)	Imagem	$0,716 \pm 0,06$	$0,617 \pm 0,02$	$0,636 \pm 0,02$	$0,662 \pm 0,03$
XGBoost	h	$0,604 \pm 0,04$	$0,705 \pm 0,04$	$0,675 \pm 0,04$	$0,650 \pm 0,04$
MLP	h	$0,686 \pm 0,03$	$0,655 \pm 0,04$	$0,662 \pm 0,04$	$0,670 \pm 0,03$
XGBoost	h e v	$0,692 \pm 0,06$	$0,780 \pm 0,06$	$0,746 \pm 0,04$	$0,731 \pm 0,04$
MLP	h e v	$0,678 \pm 0,05$	$0,669 \pm 0,06$	$0,662 \pm 0,05$	$0,668 \pm 0,03$
SCNN	t	$0,820 \pm 0,05$	$0,792 \pm 0,08$	$0,800 \pm 0,06$	$0,805 \pm 0,06$
SCNN	t e h	$0,896 \pm 0,03$	$0,800 \pm 0,03$	$0,835 \pm 0,01$	$0,845 \pm 0,01$
SCNN	t , h e v	$0,858 \pm 0,08$	$0,832 \pm 0,08$	$0,841 \pm 0,07$	$0,844 \pm 0,07$

Fonte: elaborado pelo autor.

Figura 31 – Curvas ROC para a classificação prognóstica CCLe vs. CCLp.



Fonte: elaborado pelo autor.

5.11 Análise de Significância Estatística

Para garantir que as melhorias de desempenho observadas não se devem ao acaso, foi realizada uma análise de significância estatística. O teste de McNemar (MCNEMAR, 1947), é um teste não paramétrico ideal para comparar o desempenho de dois classificadores no mesmo conjunto de dados. Assim como descrito na Seção 2.12.1, o teste foca nos casos em que as previsões dos dois modelos divergem, avaliando se um modelo é consistentemente superior ao outro.

As Tabelas de 15 a 18 apresentam os resultados da comparação entre o modelo SCNN multimodal completo (t , h e v) contra as versões com atributos de textura (t), textura e forma (t e h), o modelo de referência e o classificador XGBoost.

As abordagens baseadas em aprendizado profundo e sensíveis à assimetria (modelos SCNN) superam significativamente os métodos que dependem apenas dos atributos do SC (classificador XGBoost). Em quase todos os cenários, especialmente nos mais desafiadores (CN vs. DA e CCLe vs. CCLp), a diferença entre o modelo completo e os modelos baseados apenas em SC e/ou tecido foi altamente significativa ($p \ll 0,05$).

No entanto, a análise revela uma nuance importante: a inclusão dos atributos volumétricos e descritores SC, embora melhore as métricas de desempenho, nem sempre resulta em uma vantagem estatisticamente significativa sobre o modelo . Por exemplo, na tarefa CN vs. DA, a diferença entre o modelos SCNN com atributos de textura (t), e (t, h, v) não foi significativa ($p = 0,8070$).

Isso sugere que as características de textura do mapa de energia da DT-CWT aprendidas pelo módulo Convolucional da rede Siamesa já capturam uma porção substancial da informação preditiva contida também nos descritores de forma e volume. A fusão das três modalidades ajuda a refinar e estabilizar o modelo, levando a um melhor desempenho geral, mas a maior parte do ganho de poder discriminativo provém da própria arquitetura Siamesa e atributos de textura. Nesse sentido, esses resultados indicam uma estagnação do modelo com a adição desses novos atributos e aponta para a necessidade de explorar modalidades de dados diferentes para obter ganhos adicionais.

Tabela 15 – Resultados do Teste de McNemar para a Classificação CN vs. CCL

Comparação	χ^2	p-valor
SCNN (t , h e t) vs. SCNN (Baseline)	5,01	0,0252
SCNN (t , h e t) vs. XGBoost (h e v)	4,33	0,0374
SCNN (t , h e t) vs. XGBoost (h)	3,45	0,0634
SCNN (t , h e t) vs. SCNN (t e h)	0,31	0,5758
SCNN (t , h e t) vs. SCNN (t)	0,14	0,7115

Fonte: elaborado pelo autor.

Tabela 16 – Resultados do Teste de McNemar para a Classificação CN vs. DA

Comparação	χ^2	p-valor
SCNN (t , h e t) vs. SCNN (Baseline)	29,26	< 0,0001
SCNN (t , h e t) vs. XGBoost (h)	6,45	0,0111
SCNN (t , h e t) vs. XGBoost (h e v)	1,97	0,1602
SCNN (t , h e t) vs. SCNN (t)	0,06	0,8070
SCNN (t , h e t) vs. SCNN (t e h)	0,01	0,9035

Fonte: elaborado pelo autor.

Tabela 17 – Resultados do Teste de McNemar para a Classificação CCL vs. DA

Comparação	χ^2	p-valor
SCNN (t , h e t) vs. SCNN (Baseline)	7,20	0,0073
SCNN (t , h e t) vs. XGBoost (h)	2,06	0,1509
SCNN (t , h e t) vs. XGBoost (h e v)	1,37	0,2416
SCNN (t , h e t) vs. SCNN (t)	0,41	0,5228
SCNN (t , h e t) vs. SCNN (t e h)	0,23	0,6304

Fonte: elaborado pelo autor.

Tabela 18 – Resultados do Teste de McNemar para a Classificação CCLe vs. CCLp

Comparação	χ^2	p-valor
SCNN (t , h e t) vs. SCNN (Baseline)	28,46	< 0,0001
SCNN (t , h e t) vs. XGBoost (h)	19,38	< 0,0001
SCNN (t , h e t) vs. XGBoost (h e v)	7,48	0,0062
SCNN (t , h e t) vs. SCNN (t)	1,09	0,2955
SCNN (t , h e t) vs. SCNN (t e h)	0,77	0,3186

Fonte: elaborado pelo autor.

5.12 Considerações Finais

Em todas as tarefas de classificação, os resultados confirmam o valor agregado da inclusão de características de assimetria inter-hemisférica em arquiteturas de aprendizado profundo para avaliação da neurodegeneração. Em tarefas de estágio inicial (CN vs. CCL e CCLe vs. CCLp), assimetrias geométricas e características volumétricas fornecem informações complementares, enquanto em estágios posteriores (CN vs. DA), os padrões de textura profundos capturados pela DT-CWT são especialmente eficazes.

A dificuldade na classificação CCL vs. DA destaca as limitações das características estruturais convencionais e ressalta a importância de combinar análises de textura espacialmente localizadas e sensíveis à orientação com descritores geométricos de ordem superior. Importante ressaltar que o desempenho do modelo na tarefa CCLe vs. CCLp sugere um forte potencial para estratificação de risco precoce e monitoramento em ambientes clínicos, o que poderia apoiar a tomada de decisões terapêuticas e estratégias de acompanhamento individualizadas.

Os resultados, validados estatisticamente, confirmam o valor agregado da incorporação de atributos de assimetria inter-hemisférica em arquiteturas de aprendizado profundo para a avaliação da neurodegeneração, especialmente na tarefa clinicamente relevante de predição da progressão do CCL (CCLe vs. CCLp).

Capítulo 6

Conclusões

Este trabalho de mestrado propôs e validou uma metodologia computacional multimodal para a quantificação de assimetrias inter-hemisféricas no hipocampo, com aplicação direta no diagnóstico e, crucialmente, no prognóstico da DA. A abordagem foi fundamentada na hipótese de que a fusão de biomarcadores de textura, forma e volume, processados por uma arquitetura de aprendizado profundo sensível a diferenças estruturais, poderia superar as limitações das análises volumétricas tradicionais.

A principal contribuição desta pesquisa reside na concepção de um *pipeline* que integra DT-CWT e o descritor *shape context* em uma SCNN. Esta arquitetura foi projetada para modelar explicitamente as diferenças inter-hemisféricas, aprendendo uma representação discriminativa da assimetria.

Os resultados experimentais demonstraram a eficácia da abordagem, com destaque para a tarefa de maior relevância clínica e complexidade: a predição da conversão de CCL para DA. Nesta tarefa (CCLe vs.CCLp), a abordagem que combinou o modelo SCNN com os descritores de forma (t e h) alcançou um F1-score de 0,845, evidenciando que a fusão da assimetria textural, aprendida pela rede, com a assimetria geométrica, capturada pelo *shape context*, tem capacidade de ser um preditor da progressão da doença.

A análise comparativa revelou uma concordância entre as modalidades de atributos que é dependente do estágio da doença. Na distinção entre indivíduos CN vs. CCL, as características de forma se mostraram mais informativas, ao passo que na classificação entre CN vs. DA, a integração completa de textura, forma e volume (F1-score de 0,872) provou ser a mais robusta. Este padrão corrobora a hipótese de que diferentes aspectos da patologia (deformação geométrica, degeneração textural e atrofia volumétrica) dominam em diferentes fases da neurodegeneração.

A análise de significância estatística confirmou que a introdução da arquitetura Sia-

mesa com atributos de textura representa um avanço substancial sobre os demais modelos. Contudo, também apontou para uma parcial redundância de informação entre os atributos em estágios mais avançados da doença, destacando a relevância das características de textura baseadas em *wavelets* codificadas pela rede.

Os resultados obtidos neste trabalho corroboram e expandem as conclusões apresentadas na literatura. A eficácia da assimetria hipocampal na predição da conversão de CCL para DA alinha-se aos resultados de Long, Jiang e Zhang (2018) e Wachinger et al. (2016), que identificaram a assimetria em estruturas do sistema límbico como um biomarcador prognóstico relevante. A abordagem multimodal proposta, que integra descritores de textura e forma, representa um avanço sobre metodologias que exploraram essas características de maneira isolada, como as análises de texturas de Oliveira, Poloni e Ferrari (2020) e Poloni et al. (2021) ou as abordagens de análise geométrica, como o *BrainPrint* (WACHINGER et al., 2015). Ademais, a utilização de uma arquitetura de aprendizado profundo para modelar as diferenças inter-hemisféricas reflete a tendência atual de empregar modelos computacionais avançados para tarefas prognósticas, conforme apontado por revisões recentes (WANG et al., 2024; NIYAS; THIYAGARAJAN, 2023).

Em suma, este trabalho demonstrou com que a análise multimodal da assimetria inter-hemisférica hipocampal constitui uma ferramenta relevante para o estudo da DA. A metodologia proposta apresentou um desempenho competitivo em tarefas prognósticas desafiadoras. Somado a isso, também ofereceu uma compreensão mais granular sobre quais biomarcadores de imagem são mais pertinentes em diferentes estágios da doença.

6.1 Trabalhos Futuros

Apesar dos resultados promissores, este trabalho possui limitações e questionamentos que podem ser solucionados por pesquisas futuras. As seguintes direções são propostas:

- ❑ **Expansão da Análise Anatômica:** A análise pode ser estendida para incluir outras regiões corticais e subcorticais que também são afetadas pela DA, como a amígdala e o córtex entorrinal, para investigar padrões de assimetria em uma escala mais ampla.
- ❑ **Integração de Outras Modalidades:** Futuramente, a metodologia pode ser enriquecida com a incorporação de outras modalidades de imagem, como a ressonância magnética funcional (fMRI) e a tomografia por emissão de pósitrons (PET), ou com biomarcadores bioquímicos (ex: dados de líquido cefalorraquidiano). A fusão de dados estruturais, funcionais e moleculares tem o potencial de criar modelos ainda mais precisos.
- ❑ **Validação em Conjuntos de Dados Externos:** Para garantir a generalização dos resultados, a metodologia deve ser validada em conjuntos de dados maiores

e mais diversos, provenientes de diferentes centros de pesquisa e com diferentes protocolos de aquisição de imagem.

- ❑ **Análise Longitudinal Aprofundada:** Embora dados longitudinais tenham sido utilizados, uma análise mais aprofundada da trajetória temporal das assimetrias em cada indivíduo poderia fornecer *insights* valiosos sobre a velocidade e os padrões de progressão da doença, correlacionando as mudanças de imagem com o declínio cognitivo ao longo do tempo.
- ❑ **Exploração de Arquiteturas de Aprendizado Profundo:** A investigação de arquiteturas de redes neurais mais complexas, como modelos baseados em *Transformers* ou redes neurais de grafos para analisar as relações entre os pontos salientes, pode levar a ganhos adicionais de desempenho.

Espera-se que a continuidade desta linha de pesquisa contribua para o desenvolvimento de ferramentas de diagnóstico auxiliado por computador cada vez mais precisas e confiáveis, com o objetivo final de impactar positivamente o manejo clínico e a qualidade de vida dos pacientes com Doença de Alzheimer.

Referências

ABDI, H.; WILLIAMS, L. J. Tukey's Honestly Significant Difference (HSD) Test. **Encyclopedia of Research Design**, v. 3, p. 1–5, 2010.

AGOSTINHO, D.; SIMÕES, M.; CASTELO-BRANCO, M. Predicting conversion from mild cognitive impairment to alzheimer's disease: a multimodal approach. **Brain Communications**, Oxford University Press (OUP), v. 6, n. 4, 2024. ISSN 2632-1297. Disponível em: <<http://dx.doi.org/10.1093/braincomms/fcae208>>.

Alzheimer's Association. **Alzheimer's Association**. 2023. <<https://www.alz.org/>>. Acesso em: 27 de julho de 2023.

AVANTS, B. B. et al. An open source multivariate framework for n-tissue segmentation with evaluation on public data. **Neuroinformatics**, Springer, v. 9, n. 4, p. 381–400, 2011.

BELAROUSSI, B. et al. Intensity non-uniformity correction in MRI: existing methods and their validation. **Medical Image Analysis**, Elsevier, v. 10, n. 2, p. 234–246, 2006.

BELONGIE, S.; MALIK, J.; PUZICHA, J. Shape matching and object recognition using shape contexts. **Pattern Analysis and Machine**, v. 24, n. 4, p. 509–522, 2002.

BERGER, V. W.; ZHOU, Y. Kolmogorov–Smirnov Test: Overview. In: _____. **Wiley StatsRef: Statistics Reference Online**. [S.l.]: John Wiley & Sons, Ltd, 2014. ISBN 9781118445112.

BRASSAROTE, G.; SOUZA, E.; MONICO, J. Non-decimated Wavelet Transform for a Shift-invariant Analysis. **TEMA (São Carlos)**, Sociedade Brasileira de Matemática Aplicada e Computacional, v. 19, n. 1, p. 93–110, Jan 2018. ISSN 2179-8451. Disponível em: <<https://doi.org/10.5540/tema.2018.019.01.0093>>.

BREIJYEH, Z.; KARAMAN, R. Comprehensive Review on Alzheimer's Disease: Causes and Treatment. **Molecules**, v. 25, n. 24, p. 5789, dez. 2020. ISSN 1420-3049. Disponível em: <<https://www.mdpi.com/1420-3049/25/24/5789>>.

BROMLEY, J. et al. Signature verification using a "siamese" time delay neural network. In: COWAN, J.; TESAURO, G.; ALSPECTOR, J. (Ed.). **Advances in Neural Information Processing Systems**. Morgan-Kaufmann, 1993. v. 6. Disponível em: <https://proceedings.neurips.cc/paper_files/paper/1993/file/288cc0ff022877bd3df94bc9360b9c5d-Paper.pdf>.

BUADES, A.; COLL, B.; MOREL, J.-M. A Non-Local Algorithm for Image Denoising. In: **2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)**. San Diego, CA, USA: IEEE, 2005. v. 2, p. 60–65. ISBN 978-0-7695-2372-9. Disponível em: <<http://ieeexplore.ieee.org/document/1467423/>>.

CHAUVIN, L. et al. Efficient Pairwise Neuroimage Analysis Using the Soft Jaccard Index and 3D Keypoint Sets. **IEEE Transactions on Medical Imaging**, v. 41, n. 4, p. 836–845, 2022.

CHEN, T.; GUESTRIN, C. XGBoost: A Scalable Tree Boosting System. In: **Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining**. Cidade de Nova Iorque, Nova Iorque: Association for Computer Machinery, 2016. p. 785–794.

CHOI, P. T.; LUI, L. M. Fast Disk Conformal Parameterization of Simply-Connected Open Surfaces. **Journal of Scientific Computing**, Springer Science and Business Media LLC, v. 65, n. 3, p. 1065–1090, fev. 2015. ISSN 1573-7691. Disponível em: <<http://dx.doi.org/10.1007/s10915-015-9998-2>>.

CLAUS, J. J. et al. Practical use of visual medial temporal lobe atrophy cut-off scores in Alzheimer's disease: Validation in a large memory clinic population. **European Radiology**, v. 27, n. 8, p. 3147–3155, ago. 2017. ISSN 0938-7994, 1432-1084. Disponível em: <<http://link.springer.com/10.1007/s00330-016-4726-3>>.

CORDER, G. W.; FOREMAN, D. I. **Nonparametric statistics**. 2. ed. Nashville, TN: John Wiley & Sons, 2014.

COUPÉ, P. et al. An optimized blockwise nonlocal means denoising filter for 3D magnetic resonance images. **IEEE Transactions on Medical Imaging**, IEEE, v. 27, n. 4, p. 425–441, 2008.

COX, D. R. Regression Models and Life-Tables. **Journal of the Royal Statistical Society: Series B (Methodological)**, v. 34, n. 2, p. 187–202, 1972.

DUBOIS, B. et al. Biomarkers in Alzheimer's disease: role in early and differential diagnosis and recognition of atypical variants. **Alzheimer's Research & Therapy**, v. 15, n. 175, p. 1–13, 2023.

EDWARDS, A. L. Note on the “correction for continuity” in testing the significance of the difference between correlated proportions. **Psychometrika**, Cambridge University Press (CUP), v. 13, n. 3, p. 185–187, set. 1948. ISSN 1860-0980. Disponível em: <<http://dx.doi.org/10.1007/BF02289261>>.

ELBERT, T. et al. Increased Cortical Representation of the Fingers of the Left Hand in String Players. **Science**, v. 270, n. 5234, p. 305–307, 1995. Disponível em: <<https://www.science.org/doi/abs/10.1126/science.270.5234.305>>.

FAUQUEUR, J.; KINGSBURY, N.; ANDERSON, R. Multiscale Keypoint Detection using the Dual-Tree Complex Wavelet Transform. In: **2006 International Conference on Image Processing**. Atlanta, GA: IEEE, 2006. p. 1625–1628. ISBN 978-1-4244-0480-3. Disponível em: <<https://ieeexplore.ieee.org/document/4106857/>>.

- FERRARI, R. J.; PINTO, C. H. V.; MOREIRA, C. A. F. Detection of the midsagittal plane in MR images using a sheetness measure from eigenanalysis of local 3D phase congruency responses. In: **2016 IEEE International Conference on Image Processing (ICIP)**. [S.l.: s.n.], 2016. p. 2335–2339.
- FROME, A. et al. Recognizing Objects in Range Data Using Regional Point Descriptors. In: KANADE, T. et al. (Ed.). **Computer Vision - ECCV 2004**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004. v. 3023, p. 224–237. ISBN 978-3-540-21982-8 978-3-540-24672-5. Series Title: Lecture Notes in Computer Science. Disponível em: <http://link.springer.com/10.1007/978-3-540-24672-5_18>.
- GAUTHIER, S. et al. **World Alzheimer Report 2021**. [S.l.], 2021.
- _____. **World Alzheimer Report 2021**. [S.l.], 2021.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep learning**. [S.l.]: The MIT Press, 2016.
- GREWAL, M. et al. An end-to-end deep learning approach for landmark detection and matching in medical images. In: IŞGUM, I.; LANDMAN, B. A. (Ed.). **Medical Imaging 2020: Image Processing**. SPIE, 2020. v. 11313, p. 1131328. Disponível em: <<https://doi.org/10.1117/12.2549302>>.
- _____. Automatic landmark correspondence detection in medical images with an application to deformable image registration. **Journal of Medical Imaging**, v. 10, n. 01, fev. 2023. ISSN 2329-4302. Disponível em: <<https://www.spiedigitallibrary.org/journals/journal-of-medical-imaging/volume-10/issue-01/014007/Automatic-landmark-correspondence-detection-in-medical-images-with-an-application/10.1117/1.JMI.10.1.014007.full>>.
- GUO, Y. et al. Automatic landmark detection and registration of brain cortical surfaces via quasi-conformal geometry and convolutional neural networks. **Computers in Biology and Medicine**, v. 163, p. 107185, set. 2023. ISSN 00104825. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0010482523006509>>.
- HADSELL, R.; CHOPRA, S.; LECUN, Y. Dimensionality reduction by learning an invariant mapping. In: **2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)**. [S.l.: s.n.], 2006. v. 2, p. 1735–1742.
- HASSEN, S. B. et al. Deep learning methods for early detection of Alzheimer's disease using structural MR images: a survey. **Neurocomputing**, Elsevier BV, v. 576, p. 1–17, 2024.
- HOOPEs, A. et al. Synthstrip: skull-stripping for any brain image. **NeuroImage**, v. 260, p. 119474, 2022. ISSN 1053-8119. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1053811922005900>>.
- HOU, Y. et al. Ageing as a risk factor for neurodegenerative disease. **Nature Reviews Neurology**, v. 15, n. 10, p. 565–581, out. 2019. ISSN 1759-4758, 1759-4766. Disponível em: <<https://www.nature.com/articles/s41582-019-0244-7>>.
- IZENMAN, A. J. Linear discriminant analysis. In: _____. **Modern Multivariate Statistical Techniques**. Springer New York, 2013. p. 237–280. ISBN 9780387781891. Disponível em: <http://dx.doi.org/10.1007/978-0-387-78189-1_8>.

JACK, C. R. et al. Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. **Alzheimer's & Dementia**, v. 7, n. 3, p. 257–262, maio 2011. ISSN 1552-5260, 1552-5279. Disponível em: <<https://alz-journals.onlinelibrary.wiley.com/doi/10.1016/j.jalz.2011.03.004>>.

_____. Revised criteria for diagnosis and staging of Alzheimer's disease: Alzheimer's Association Workgroup. **Alzheimer's & Dementia**, v. 20, n. 8, p. 5143–5169, ago. 2024. ISSN 1552-5260, 1552-5279. Disponível em: <<https://alz-journals.onlinelibrary.wiley.com/doi/10.1002/alz.13859>>.

JONKER, R.; VOLGENANT, A. A shortest augmenting path algorithm for dense and sparse linear assignment problems. **Computing**, v. 38, p. 325–340, 1987.

KINGSBURY, N. G. The dual-tree complex wavelet transform: a new technique for shift invariance and directional filters. In: CITESEER. **IEEE digital signal processing workshop**. [S.l.], 1998. v. 86, p. 120–131.

KLEIN, S. et al. Elastix: A Toolbox for Intensity-Based Medical Image Registration. **IEEE Transactions on Medical Imaging**, v. 29, n. 1, p. 196–205, 2009.

_____. elastix: a toolbox for intensity-based medical image registration. **IEEE Transactions on Medical Imaging**, v. 29, n. 1, p. 196 – 205, January 2010.

KNOPMAN, D. et al. Alzheimer's disease. **Nature Reviews Disease Primers**, Springer Nature, v. 7, n. 33, 2021.

KOCH, G. R. Siamese neural networks for one-shot image recognition. In: . [s.n.], 2015. Disponível em: <<https://api.semanticscholar.org/CorpusID:13874643>>.

KONG, X. et al. Mapping brain asymmetry in health and disease through the ENIGMA consortium. **Human Brain Mapping**, v. 43, n. 1, p. 167–181, jan. 2022. ISSN 1065-9471, 1097-0193. Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1002/hbm.25033>>.

KONG, X.-Z. et al. Mapping cortical brain asymmetry in 17,141 healthy individuals worldwide via the ENIGMA Consortium. **Proceedings of the National Academy of Sciences**, v. 115, n. 22, p. E5154–E5163, 2018. Disponível em: <<https://www.pnas.org/doi/abs/10.1073/pnas.1718418115>>.

KORB, M. M.; FERRARI, R. J. Automatic Positioning of Hippocampus Deformable Mesh Models in Brain MR Images Using a Weighted 3D-SIFT Technique. In: GERVASI, O. et al. (Ed.). **Computational Science and Its Applications – ICCSA 2020**. Cham: Springer International Publishing, 2020. p. 75–90. ISBN 978-3-030-58802-1.

KORBMACHER, M. et al. Distinct longitudinal brain white matter microstructure changes and associated polygenic risk of common psychiatric disorders and alzheimer's disease in the uk biobank. **Biological Psychiatry Global Open Science**, Elsevier BV, v. 4, n. 4, p. 100323, jul. 2024. ISSN 2667-1743. Disponível em: <<http://dx.doi.org/10.1016/j.bpsgos.2024.100323>>.

- KUO, F.; MASSOUD, T. F. Structural asymmetries in normal brain anatomy: A brief overview. **Annals of Anatomy - Anatomischer Anzeiger**, v. 241, p. 1–10, abr. 2022. ISSN 09409602. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0940960222000097>>.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, Springer Science and Business Media LLC, v. 521, n. 7553, p. 436–444, maio 2015. ISSN 1476-4687. Disponível em: <<http://dx.doi.org/10.1038/nature14539>>.
- LEE, H.-S. et al. Speaker verification using kernel-based binary classifiers with binary operation derived features. In: **2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)**. [S.l.: s.n.], 2014. p. 1660–1664.
- LI, G.; JI, Z.; SUN, Q. Deep multi-instance conv-transformer frameworks for landmark-based brain mri classification. **Electronics**, MDPI AG, v. 13, n. 5, p. 980, mar. 2024. ISSN 2079-9292. Disponível em: <<http://dx.doi.org/10.3390/electronics13050980>>.
- LIU, C.; XU, J.; WANG, F. A Review of Keypoints' Detection and Feature Description in Image Registration. **Scientific Programming**, v. 2021, p. 1–25, dez. 2021. ISSN 1875-919X, 1058-9244. Disponível em: <<https://www.hindawi.com/journals/sp/2021/8509164/>>.
- LIU, J. et al. Attention-guided 3d cnn with lesion feature selection for early alzheimer's disease prediction using longitudinal smri. **IEEE Journal of Biomedical and Health Informatics**, v. 29, n. 1, p. 324–332, 2025.
- LONG, X.; JIANG, C.; ZHANG, L. Morphological Biomarker Differentiating MCI Converters from Nonconverters: Longitudinal Evidence Based on Hemispheric Asymmetry. **Behavioural Neurology**, v. 2018, p. 1–9, 2018. ISSN 0953-4180, 1875-8584. Disponível em: <<https://www.hindawi.com/journals/bn/2018/3954101/>>.
- LONG, X. et al. Distinct laterality alterations distinguish mild cognitive impairment and Alzheimer's disease from healthy aging: Statistical parametric mapping with high resolution MRI. **Human Brain Mapping**, v. 34, n. 12, p. 3400–3410, dez. 2013. ISSN 1065-9471, 1097-0193. Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1002/hbm.22157>>.
- LORETO, F. et al. Visual atrophy rating scales and amyloid PET status in an Alzheimer's disease clinical cohort. **Annals of Clinical and Translational Neurology**, v. 10, n. 4, p. 619–631, abr. 2023. ISSN 2328-9503, 2328-9503. Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1002/acn3.51749>>.
- LOWE, D. G. Distinctive Image Features from Scale-Invariant Keypoints. **International Journal of Computer Vision**, v. 60, n. 2, p. 91–110, nov. 2004. ISSN 0920-5691. Disponível em: <<http://link.springer.com/10.1023/B:VISI.0000029664.99615.94>>.
- LUBBEN, N. et al. The enigma and implications of brain hemispheric asymmetry in neurodegenerative diseases. **Brain Communications**, v. 3, n. 3, p. fcab211, jul. 2021. ISSN 2632-1297. Disponível em: <<https://academic.oup.com/braincomms/article/doi/10.1093/braincomms/fcab211/6364895>>.

LUO, M. et al. Class activation attention transfer neural networks for MCI conversion prediction. **Computers in Biology and Medicine**, v. 156, p. 106700, 2023.

_____. Dual attention based fusion network for mci conversion prediction. **Computers in Biology and Medicine**, Elsevier BV, v. 182, p. 109039, nov. 2024. ISSN 0010-4825. Disponível em: <<http://dx.doi.org/10.1016/j.compbimed.2024.109039>>.

MAIER, A. et al. A gentle introduction to deep learning in medical image processing. **Zeitschrift für Medizinische Physik**, Elsevier BV, v. 29, n. 2, p. 86–101, maio 2019. ISSN 0939-3889. Disponível em: <<http://dx.doi.org/10.1016/j.zemedi.2018.12.003>>.

MATSUDA, S. et al. Brain structural asymmetry associated with cognitive impairment in community-dwelling older adults. **Archives of Gerontology and Geriatrics**, Elsevier BV, v. 138, p. 105975, nov. 2025. ISSN 0167-4943. Disponível em: <<http://dx.doi.org/10.1016/j.archger.2025.105975>>.

MAZZOLA, A. A. Ressonância magnética: princípios de formação da imagem e aplicações em imagem funcional. **Revista Brasileira de Física Médica**, v. 3, n. 1, p. 117–129, out. 2009. Disponível em: <<https://www.rbfm.org.br/rbfm/article/view/51>>.

MCKHANN, G. et al. Clinical diagnosis of Alzheimer's disease: Report of the NINCDS—ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. **Neurology**, v. 77, n. 4, p. 333–333, jul. 1984. ISSN 0028-3878, 1526-632X. Disponível em: <<https://www.neurology.org/doi/10.1212/01.wnl.0000400650.92875.cf>>.

MCNEMAR, Q. Note on the sampling error of the difference between correlated proportions or percentages. **Psychometrika**, v. 12, n. 2, p. 153–157, 1947.

MISHRA, P. et al. Application of student's t-test, analysis of variance, and covariance. **Annals of Cardiac Anaesthesia**, v. 22, n. 4, p. 407–411, 2019. ISSN 0971-9784.

MONTEMBEAULT, M. et al. Clinical, anatomical, and pathological features in the three variants of primary progressive aphasia: A review. **Frontiers in Neurology**, v. 9, 2018. ISSN 1664-2295. Disponível em: <<https://www.frontiersin.org/articles/10.3389/fneur.2018.00692>>.

NAFORNITA, C.; ISAR, A.; NAFORNITA, I. The Hyperanalytic Wavelet Packets - A solution to increase the directional selectivity in image analysis. In: **2012 10th International Symposium on Electronics and Telecommunications**. Timisoara, Timis, Romania: IEEE, 2012. p. 231–234. ISBN 978-1-4673-1176-2 978-1-4673-1177-9 978-1-4673-1175-5. Disponível em: <<http://ieeexplore.ieee.org/document/6408041/>>.

NEMATİ, S. S. et al. Lateralization of the hippocampus: A review of molecular, functional, and physiological properties in health and disease. **Behavioural Brain Research**, v. 454, p. 114657, out. 2023. ISSN 01664328. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0166432823003753>>.

NIYAS, K. M.; THIYAGARAJAN, P. A Systematic Review on Early Prediction of Mild Cognitive Impairment to Alzheimers Using Machine Learning Algorithms. **International Journal of Intelligent Networks**, v. 4, p. 74–88, 2023.

NYÚL, L. G.; UDUPA, J. K.; ZHANG, X. New variants of a method of MRI scale standardization. **IEEE Transactions on Medical Imaging**, IEEE, v. 19, n. 2, p. 143–150, 2000.

OCKLENBURG, S. et al. Clinical implications of brain asymmetries. **Nature Reviews Neurology**, Springer Science and Business Media LLC, v. 20, n. 7, p. 383–394, maio 2024. ISSN 1759-4766. Disponível em: <<http://dx.doi.org/10.1038/s41582-024-00974-8>>.

OLIVEIRA, I. A. Duarte de; POLONI, K. M.; FERRARI, R. J. Exploring Hippocampal Asymmetrical Features from Magnetic Resonance Images for the Classification of Alzheimer's Disease. In: **2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS)**. [S.l.: s.n.], 2020. p. 59–64.

OOSTVEEN, W. M. V.; LANGE, E. C. M. D. Imaging Techniques in Alzheimer's Disease: A Review of Applications in Early Diagnosis and Longitudinal Monitoring. **International Journal of Molecular Sciences**, v. 22, n. 4, p. 2110, fev. 2021. ISSN 1422-0067. Disponível em: <<https://www.mdpi.com/1422-0067/22/4/2110>>.

POLONI, K. et al. Brain MR image classification for Alzheimer's disease diagnosis using structural hippocampal asymmetrical attributes from directional 3-D log-Gabor filter responses. **Neurocomputing**, v. 419, p. 126–135, 2021.

PRINCE, M. et al. **World Alzheimer Report 2015**. [S.l.], 2015.

RAHIM, N. et al. Early progression detection from mci to ad using multi-view mri for enhanced assisted living. **Image and Vision Computing**, Elsevier BV, v. 157, p. 105491, maio 2025. ISSN 0262-8856. Disponível em: <<http://dx.doi.org/10.1016/j.imavis.2025.105491>>.

RASMUSSEN, J.; LANGERMAN, H. Alzheimer's Disease - Why We Need Early Diagnosis. **Degenerative Neurological and Neuromuscular Disease**, Volume 9, p. 123–130, dez. 2019. ISSN 1179-9900. Disponível em: <<https://www.dovepress.com/alzheimers-disease-why-we-need-early-diagnosis-peer-reviewed-article-DNND>>.

RAU, A.; URBACH, H. The MTA score—simple and reliable, the best for now? **European Radiology**, v. 31, n. 12, p. 9057–9059, dez. 2021. ISSN 0938-7994, 1432-1084. Disponível em: <<https://link.springer.com/10.1007/s00330-021-08340-8>>.

REINHOLD, J. C. et al. Evaluating the impact of intensity normalization on MR image synthesis. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. **Medical Imaging 2019: Image Processing**. [S.l.], 2019. v. 10949, p. 109493H.

REUTER, M.; FISCHL, B. Avoiding asymmetry-induced bias in longitudinal image processing. **NeuroImage**, v. 57, n. 1, p. 19–21, jul. 2011. ISSN 10538119. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S1053811911002539>>.

RISTER, B.; HOROWITZ, M. A.; RUBIN, D. L. Volumetric Image Registration From Invariant Keypoints. **IEEE Transactions on Image Processing**, v. 26, n. 10, p. 4900–4910, out. 2017. ISSN 1057-7149, 1941-0042. Disponível em: <<http://ieeexplore.ieee.org/document/7967757/>>.

- RONNEBERGER, O.; FISCHER, P.; BROX, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: NAVAB, N. et al. (Ed.). **Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015**. Cham: Springer International Publishing, 2015. p. 234–241. ISBN 978-3-319-24574-4.
- ROY, S. K. et al. Siamese networks: The tale of two manifolds. In: **Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)**. [S.l.: s.n.], 2019.
- SAHINER, B. et al. Deep learning in medical imaging and radiation therapy. **Medical Physics**, Wiley, v. 46, n. 1, nov. 2018. ISSN 2473-4209. Disponível em: <<http://dx.doi.org/10.1002/mp.13264>>.
- SCHELTENS, P. et al. Atrophy of medial temporal lobes on MRI in probable Alzheimer's disease and normal ageing: diagnostic value and neuropsychological correlates. **Journal of Neurology, Neurosurgery & Psychiatry**, BMJ Publishing Group Ltd, v. 55, n. 10, p. 967–972, 1992. ISSN 0022-3050. Disponível em: <<https://jnnp.bmj.com/content/55/10/967>>.
- SCHILLING, L. P. et al. Diagnóstico da doença de Alzheimer: recomendações do Departamento Científico de Neurologia Cognitiva e do Envelhecimento da Academia Brasileira de Neurologia. **Dementia & Neuropsychologia**, Academia Brasileira de Neurologia, Departamento de Neurologia Cognitiva e Envelhecimento, v. 16, n. 3, p. 25–39, Sep 2022. ISSN 1980-5764.
- SELESNICK, I.; BARANIUK, R.; KINGSBURY, N. The dual-tree complex wavelet transform. **IEEE Signal Processing Magazine**, v. 22, n. 6, p. 123–151, nov. 2005. ISSN 1053-5888. Disponível em: <<http://ieeexplore.ieee.org/document/1550194/>>.
- SELESNICK, I. W.; LI, K. Y. Video denoising using 2D and 3D dual-tree complex wavelet transforms. In: UNSER, M. A.; ALDROUBI, A.; LAINE, A. F. (Ed.). San Diego, California, USA: [s.n.], 2003. p. 607.
- SHAMONIN, D. P. et al. Fast Parallel Image Registration on CPU and GPU for Diagnostic Classification of Alzheimer's Disease. **Frontiers in Neuroinformatics**, v. 7, n. 50, p. 1–15, January 2014.
- SINGH, A.; KINGSBURY, N. Efficient Convolutional Network Learning Using Parametric Log Based Dual-Tree Wavelet ScatterNet. In: **2017 IEEE International Conference on Computer Vision Workshops (ICCVW)**. [S.l.]: IEEE, 2017. p. 1140–1147.
- SLED, J.; ZIJDENBOS, A.; EVANS, A. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. **IEEE Transactions on Medical Imaging**, v. 17, n. 1, p. 87–97, 1998.
- SOKOLOVA, M.; LAPALME, G. A systematic analysis of performance measures for classification tasks. **Information Processing and Management**, Elsevier BV, v. 45, n. 4, p. 427–437, jul. 2009. ISSN 0306-4573. Disponível em: <<http://dx.doi.org/10.1016/j.ipm.2009.03.002>>.

- SOUZA, B. D. S.; POLONI, K. M.; FERRARI, R. J. Detector of 3-D salient points based on the dual-tree complex wavelet transform for the positioning of hippocampi meshes in magnetic resonance images. **Journal of Neuroscience Methods**, v. 341, p. 108789, jul. 2020. ISSN 01650270. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0165027020302120>>.
- SRIVASTAVA, N. et al. Dropout: A simple way to prevent neural networks from overfitting. **Journal of Machine Learning Research**, v. 15, n. 56, p. 1929–1958, 2014. Disponível em: <<http://jmlr.org/papers/v15/srivastava14a.html>>.
- TAKIZAWA, C. et al. Epidemiological and Economic Burden of Alzheimer’s Disease: A Systematic Literature Review of Data across Europe and the United States of America. **Journal of Alzheimer’s Disease**, IOS Press, v. 43, n. 4, p. 1271–1284, 2015.
- TOEWS, M.; WELLS, W. M. Efficient and robust model-to-image alignment using 3D scale-invariant features. **Medical Image Analysis**, v. 17, n. 3, p. 271–282, 2013. ISSN 1361-8415. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1361841512001533>>.
- TUSTISON, N. J. et al. N4ITK: Improved N3 Bias Correction. **IEEE Transactions on Medical Imaging**, v. 29, n. 6, p. 1310–1320, 2010.
- Villa Pinto, C. H.; FERRARI, R. J. Initialization of deformable models in 3D magnetic resonance images guided by automatically detected phase congruency point landmarks. **Pattern Recognition Letters**, v. 79, p. 1–7, 2016. ISSN 0167-8655.
- VINCENT, P.; BENGIO, Y. Manifold Parzen Windows. In: BECKER, S.; THRUN, S.; OBERMAYER, K. (Ed.). **Advances in Neural Information Processing Systems**. MIT Press, 2002. v. 15. Disponível em: <https://proceedings.neurips.cc/paper_files/paper/2002/file/2d969e2cee8cfa07ce7ca0bb13c7a36d-Paper.pdf>.
- VOVK, U.; PERNUS, F.; LIKAR, B. A review of methods for correction of intensity inhomogeneity in MRI. **IEEE Transactions on Medical Imaging**, IEEE, v. 26, n. 3, p. 405–421, 2007.
- WACHINGER, C. et al. BrainPrint: A discriminative characterization of brain morphology. **NeuroImage**, v. 109, p. 232–248, abr. 2015. ISSN 10538119. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S1053811915000476>>.
- _____. Whole-brain analysis reveals increased neuroanatomical asymmetries in dementia for hippocampus and amygdala. **Brain**, v. 139, n. 12, p. 3253–3266, dez. 2016. ISSN 0006-8950, 1460-2156. Disponível em: <<https://academic.oup.com/brain/article-lookup/doi/10.1093/brain/aww243>>.
- WANG, H.; YUSHKEVICH, P. A. Multi-atlas segmentation with joint label fusion and corrective learning—an open source implementation. **Frontiers in Neuroinformatics**, Frontiers Media SA, v. 7, 2013. ISSN 1662-5196. Disponível em: <<http://dx.doi.org/10.3389/fninf.2013.00027>>.
- WANG, X. et al. Predictive models of alzheimer’s disease dementia risk in older adults with mild cognitive impairment: a systematic review and critical appraisal. **BMC Geriatrics**, Springer Science and Business Media LLC, v. 24, n. 1, jun. 2024. ISSN 1471-2318. Disponível em: <<http://dx.doi.org/10.1186/s12877-024-05044-8>>.

World Health Organization. **Risk reduction of cognitive decline and dementia: WHO guidelines**. Geneva: World Health Organization, 2019. Section: xiii, 78 p. ISBN 978-92-4-155054-3. Disponível em: <<https://iris.who.int/handle/10665/312180>>.

WöRZ, S.; ROHR, K. Localization of anatomical point landmarks in 3D medical images by fitting 3D parametric intensity models. **Medical Image Analysis**, v. 10, n. 1, p. 41–58, fev. 2006. ISSN 13618415. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S136184150500037X>>.

YOUSAF, T.; DERVENOULAS, G.; POLITIS, M. Advances in MRI Methodology. In: **International Review of Neurobiology**. Elsevier, 2018. v. 141, p. 31–76. ISBN 978-0-12-815418-2. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0074774218300758>>.