

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

**Vivian Ferreira Margarido**

**Predição de Casos de  
Comprometimento Cognitivo Leve  
Estável e Progressivo Usando Atributos  
de Imagens de Ressonância Magnética  
para Auxílio no Diagnóstico da Doença  
de Alzheimer**

São Carlos - SP  
Julho de 2025

**Vivian Ferreira Margarido**

**Predição de Casos de  
Comprometimento Cognitivo Leve  
Estável e Progressivo Usando Atributos  
de Imagens de Ressonância Magnética  
para Auxílio no Diagnóstico 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: Visão Computacional

Orientador: Prof. Dr. Ricardo José Ferrari

São Carlos - SP

Julho de 2025



---

# Agradecimentos

---

Gostaria de expressar minha sincera gratidão a todos que contribuíram direta e indiretamente para a realização deste trabalho. Primeiramente, agradeço ao meu orientador, Prof. Dr. Ricardo José Ferrari, por sua confiança em mim, por sua constante disponibilidade e por sempre encontrar tempo para discutir dúvidas e oferecer orientações. Agradeço também aos meus colegas do laboratório por toda a colaboração e suporte.

Por fim, sou grata à minha família, ao meu namorado e aos meus amigos pelo amor, paciência e incentivo contínuo durante esta jornada acadêmica. Este trabalho é fruto de uma caminhada que não teria sido possível sem o apoio de todos vocês.



---

# Resumo

---

A Doença de Alzheimer (DA) é uma enfermidade cerebral progressiva e irreversível que afeta as funções cognitivas, sendo responsável por 60 a 80% dos casos de demência em todo o mundo. Com cerca de 50 milhões de pessoas acometidas, a DA é considerada um grave problema de saúde pública. Embora não haja cura, tratamentos podem retardar sua progressão, especialmente quando aplicados em fases precoces. Neste contexto, a predição de casos de Comprometimento Cognitivo Leve (CCL) em suas formas estável e progressiva é fundamental para intervenções mais eficazes. Este estudo tem como objetivo desenvolver um modelo preditivo baseado em Floresta Aleatória utilizando características extraídas de imagens de Ressonância Magnética (RM), como espessura cortical, volume e textura de diferentes regiões cerebrais, para auxiliar no diagnóstico precoce da DA. O método de harmonização ComBat foi aplicado para garantir a reprodutibilidade do modelo. O desempenho do modelo foi avaliado utilizando métricas como acurácia, acurácia balanceada, sensibilidade, especificidade, Área Sob a Curva (AUC) ROC e F1-score, permitindo uma análise detalhada da eficácia na previsão da progressão da DA com base em exames de RM. O melhor desempenho de F1-score foi alcançado utilizando seleção de características com base na importância da Floresta Aleatória, com o percentil 75 como ponto de corte. Um total de 312 atributos, distribuídos em 66 regiões de interesse, foi selecionado. Os resultados foram: F1-score de  $84,31 \pm 8,73\%$ , acurácia de  $84,3 \pm 8,70\%$ , acurácia balanceada de  $84,31 \pm 8,64\%$ , sensibilidade de  $84,00 \pm 9,67\%$ , especificidade de  $84,62 \pm 13,01\%$  e AUC de  $0,85 \pm 0,07$ . Os resultados demonstram que a abordagem proposta combina desempenho competitivo e interpretabilidade, oferecendo uma ferramenta promissora e clinicamente relevante para a predição da conversão de CCL para DA a partir de imagens de RM.

**Palavras-chave:** Doença de Alzheimer. CCL progressivo. CCL estável. Imagens de RM. Assinatura radiômica. Modelos preditivos.



---

# Abstract

---

Alzheimer's Disease (AD) is a progressive and irreversible brain disorder that affects cognitive functions, accounting for 60 to 80% of dementia cases worldwide. With approximately 50 million people affected, AD is considered a major public health issue. Although there is no cure, treatments can slow its progression, especially when applied in the early stages. In this context, the prediction of Mild Cognitive Impairment (MCI) cases in its stable and progressive forms is crucial for more effective interventions. This study aims to develop a predictive model based on the Random Forest algorithm using features extracted from magnetic resonance imaging (MRI), such as cortical thickness, volume, and texture from different brain regions, to aid in the early diagnosis of AD. The ComBat harmonization method was applied to ensure model reproducibility. Model performance was assessed using metrics such as accuracy, balanced accuracy, sensitivity, specificity, Area Under the Curve (AUC) and F1-score, allowing for a detailed analysis of their effectiveness in predicting AD progression based on MRI scans. The best F1-score performance was achieved using feature selection based on Random Forest importance, with the 75th percentile as the cutoff. A total of 312 features across 66 Regions of Interest (ROIs) were selected. The results were: F1-score  $84.31 \pm 8.73\%$ , accuracy  $84.31 \pm 8.70\%$ , balanced accuracy  $84.31 \pm 8.64\%$ , sensitivity  $84.00 \pm 9.67\%$ , specificity  $84.62 \pm 13.01\%$ , and AUC  $0.85 \pm 0.07$ . The results demonstrate that the proposed approach combines competitive performance and interpretability, offering a promising and clinically relevant tool for predicting the conversion from MCI to AD based on MRI.

**Keywords:** Alzheimer's Disease. Progressive MCI. Stable MCI. MRI. Radiomic signature. Predictive models.

---

# Lista de ilustrações

---

Figura 1 – Estágios da progressão da DA. . . . .	28
Figura 2 – Proteína tau em neurônios saudáveis e em DA. . . . .	30
Figura 3 – Corte coronal para visualização da substância cinzenta, substância branca e Líquido Cefalorraquidiano. . . . .	31
Figura 4 – Lobos cerebrais e suas funções. . . . .	32
Figura 5 – Comparação entre um cérebro saudável e com DA. . . . .	32
Figura 6 – <i>Spin</i> e momento magnético ( $\mu$ ) do próton de hidrogênio. . . . .	35
Figura 7 – Alinhamento dos prótons de hidrogênio em um campo magnético externo forte e o momento magnético resultante ( $M_0$ ). . . . .	36
Figura 8 – Movimento de precessão do próton de hidrogênio. . . . .	37
Figura 9 – Relaxação longitudinal e transversal. . . . .	38
Figura 10 – Sequência de pulso <i>spin</i> eco. . . . .	39
Figura 11 – Comparação entre imagens ponderadas em T1, T2 e FLAIR. . . . .	40
Figura 12 – Configuração das bobinas de gradiente. . . . .	41
Figura 13 – Efeito de aplicação de um gradiente de campo magnético na direção do eixo $z$ com amplitude 45 mT/m. . . . .	42
Figura 14 – Espaço K e a imagem de RM correspondente após a aplicação da TIF bidimensional. . . . .	43
Figura 15 – Representação esquemática da GLCM 3D. . . . .	49
Figura 16 – Exemplo da construção da matriz GLCM a partir de uma imagem 2D. . . . .	50
Figura 17 – Árvore de decisão e as regiões de decisão no espaço de objetos. . . . .	55
Figura 18 – Fluxograma das principais etapas realizadas neste projeto de pesquisa. . . . .	82
Figura 19 – Evolução longitudinal da base de dados da ADNI. . . . .	83
Figura 20 – Estratificação por idade e sexo dos pacientes CCLe da base de dados ADNI. . . . .	85
Figura 21 – Estratificação por idade e sexo dos pacientes CCLp da base de dados ADNI. . . . .	86

Figura 22 – Estratificação das imagens da base de dados ADNI por intensidade de campo magnético de cada grupo diagnóstico. . . . .	87
Figura 23 – Estratificação das imagens da base de dados ADNI por intensidade de campo magnético de cada grupo diagnóstico (sem <i>outliers</i> ). . . . .	87
Figura 24 – Etapas de pré-processamento. . . . .	88
Figura 25 – <i>Framework</i> de registro de imagens. . . . .	92
Figura 26 – Resultado da remoção do crânio em uma das imagens selecionadas da ADNI. . . . .	104
Figura 27 – Resultado da redução de ruído em uma das imagens selecionadas da ADNI. . . . .	105
Figura 28 – Resultado da correção de <i>bias field</i> em uma das imagens selecionadas da ADNI. . . . .	105
Figura 29 – Resultado da padronização de histograma de uma das imagens selecionadas da ADNI com a imagem de referência da NAC. . . . .	106
Figura 30 – Resultado do correregistro da imagem de referência da NAC com uma imagem da ADNI. . . . .	106
Figura 31 – Resultado da segmentação de tecidos cerebrais em uma imagem da ADNI normalizada com e sem padronização de histogramas. . . . .	107
Figura 32 – Resultado da segmentação das regiões de interesse em uma imagem da ADNI (vistas axial, sagital e coronal). . . . .	108
Figura 33 – Distribuição do volume total intracraniano entre as classes CCLe e CCLp (sem harmonização). . . . .	109
Figura 34 – Distribuição do volume total intracraniano entre as classes CCLe e CCLp (com harmonização). . . . .	110
Figura 35 – Distribuição do volume de cada ROI nas classes CCLe e CCLp (sem harmonização). . . . .	111
Figura 36 – Volumes médios de CSF, GM e WM em cada ROI para as classes CCLe (primeira barra) e CCLp (segunda barra) (sem harmonização). . . . .	113
Figura 37 – Volumes médios de CSF, GM e WM em cada ROI para as classes CCLe (primeira barra) e CCLp (segunda barra) (com harmonização). . . . .	114
Figura 38 – Resultado da geração da imagem de espessura cortical a partir de uma imagem da ADNI (vistas axial, sagital e coronal). . . . .	115
Figura 39 – Distribuição da espessura cortical média total entre as classes CCLe e CCLp (sem harmonização). . . . .	116
Figura 40 – Distribuição da espessura cortical média total entre as classes CCLe e CCLp (com harmonização). . . . .	117
Figura 41 – Distribuição da espessura cortical média de cada ROI nas classes CCLe e CCLp (sem harmonização). . . . .	118

Figura 42 – Distribuição da espessura cortical média de cada ROI nas classes CCLe e CCLp (com harmonização). . . . .	120
Figura 43 – Quantidade de atributos estatisticamente significativos para cada tipo de imagem e quantização antes e após harmonização ComBat. . . . .	125
Figura 44 – Efeito da harmonização ComBat sobre os atributos das classes CCLe e CCLp. . . . .	127
Figura 45 – Efeito da harmonização ComBat sobre os conjuntos de atributos da classe CCLe. . . . .	128
Figura 46 – Efeito da harmonização ComBat sobre os conjuntos de atributos da classe CCLp. . . . .	129
Figura 47 – Relação entre F1-score e quantidade de atributos selecionados. . . . .	131
Figura 48 – Importância global e contribuições locais dos 20 atributos mais importantes do modelo do Experimento 19 para previsão de CCLp. . . . .	133
Figura 49 – Interação entre os volumes de GM e WM da amígdala esquerda na previsão de CCLp pelo modelo do Experimento 19. . . . .	134

---

# Lista de tabelas

---

Tabela 1 – Sintomas da DA. . . . .	29
Tabela 2 – Domínios cognitivos que podem ser afetados pela demência. . . . .	33
Tabela 3 – Exames de avaliação do comprometimento cognitivo. . . . .	34
Tabela 4 – Diferentes ponderações para imagens de RM. . . . .	39
Tabela 5 – Hiperparâmetros e parâmetros de modelos de Floresta Aleatória. . . . .	58
Tabela 6 – Principais resultados do trabalho de Calandrelli et al. (2022). . . . .	63
Tabela 7 – Principais resultados do trabalho de Xie et al. (2023) para os grupos CCL BetaA– e CCL BetaA+. . . . .	65
Tabela 8 – Principais resultados do trabalho de Minhas et al. (2021) para distinguir CCLp e CCLe. . . . .	68
Tabela 9 – Trabalhos correlatos baseados em engenharia de atributos. . . . .	72
Tabela 10 – Trabalhos correlatos baseados em aprendizagem profunda. . . . .	78
Tabela 11 – Pontuações da Escala de Memória de Wechsler utilizadas pela ADNI para classificação dos pacientes em CCLI e CCLT. . . . .	84
Tabela 12 – Conjunto de imagens da base de dados ADNI usadas neste trabalho. . . . .	84
Tabela 13 – Conjunto de imagens da base de dados ADNI usadas neste trabalho (após remoção de outliers). . . . .	86
Tabela 14 – Regiões segmentadas utilizadas nesta proposta de pesquisa. . . . .	94
Tabela 15 – Conjuntos de atributos extraídos e suas quantidades. . . . .	109
Tabela 16 – ROIs com volumes que apresentaram diferenças estatisticamente significativas entre CCLe e CCLp (com harmonização). . . . .	112
Tabela 17 – Atributos de textura da imagem toda para cada tipo de imagem e nível de quantização que apresentaram diferenças estatisticamente significativas entre CCLe e CCLp (marcados com X) (sem harmonização). . . . .	122
Tabela 18 – ROIs que tiveram algum atributo de textura com diferença estatisticamente significativa em todos os tipos de imagens testados (sem harmonização). . . . .	124

Tabela 19 – Desempenho dos modelos de Floresta Aleatória sob diferentes configurações de pré-processamento de dados. . . . .	129
Tabela 20 – Abreviações de atributos e descrições correspondentes. . . . .	146

---

# Lista de siglas

---

<b>2D</b>	Bidimensionais
<b>3D</b>	Tridimensionais
<b>AB</b>	Áreas de Brodmann
<b>aCCLe</b>	Comprometimento Cognitivo Leve Amnésico estável
<b>aCCLp</b>	Comprometimento Cognitivo Leve Amnésico progressivo
<b>AD-NET</b>	Rede Neural Ajustada para Idade
<b>ADAS-Cog</b>	Alzheimer's Disease Assessment Scale - Cognitive Subscale
<b>ADLS</b>	Alzheimer's Disease Likelihood Score
<b>ADNI</b>	Alzheimer's Disease Neuroimaging Initiative
<b>AIC</b>	Critério de Informação de Akaike
<b>ANTs</b>	Advanced Normalization Tools
<b>ASM</b>	Momento Angular de Segunda Ordem
<b>AUC</b>	Área sob a curva ROC
<b>BE</b>	Bayesiano Empírico
<b>BetaA</b>	Beta-amiloide
<b>CAAT</b>	Transferência de Atenção por Ativação de Classe
<b>CART</b>	Classification and Regression Trees
<b>CBAM</b>	Convolutional Block Attention Module

<b>CCL</b>	Comprometimento Cognitivo Leve
<b>CCLe</b>	Comprometimento Cognitivo Leve estável
<b>CCLI</b>	Comprometimento Cognitivo Leve Inicial
<b>CCLIp</b>	Comprometimento Cognitivo Leve Inicial progressivo
<b>CCLp</b>	Comprometimento Cognitivo Leve progressivo
<b>CCLT</b>	Comprometimento Cognitivo Leve Tardio
<b>CCLTp</b>	Comprometimento Cognitivo Leve Tardio progressivo
<b>CDR</b>	Clinical Dementia Rating
<b>CDR-SOB</b>	Clinical Dementia Rating - Sum of Boxes
<b>CER</b>	Córtex Entorrinal
<b>CSF</b>	Líquido Cefalorraquidiano
<b>DA</b>	Doença de Alzheimer
<b>DCS</b>	Declínio Cognitivo Subjetivo
<b>DP</b>	Densidade de Prótons
<b>FI</b>	Fronto-insular
<b>FLAIR</b>	Fluid-Attenuated Inversion Recovery
<b>FOV</b>	Campo de Visão
<b>GLCM</b>	Matriz de Coocorrência de Níveis de Cinza
<b>GLDM</b>	Matriz de Dependência de Níveis de Cinza
<b>GLRLM</b>	Matriz de Comprimento de Corrida de Níveis de Cinza
<b>GLSZM</b>	Matriz de Zona de Tamanho de Nível de Cinza
<b>GM</b>	Substância Cinzenta
<b>GMM</b>	Modelo de Mistura de Gaussianas
<b>ICV</b>	Volume Intracraniano
<b>ID</b>	Inverse Difference
<b>IDA</b>	Image & Data Archive

<b>IDM</b>	Inverse Difference Moment
<b>IDMN</b>	Inverse Difference Moment Normalized
<b>IDN</b>	Inverse Difference Normalized
<b>ITCom</b>	Índice de Texturas Combinadas
<b>ITH</b>	Índice Textura Hipocampal
<b>IVCom</b>	Índice de Volumes Combinados
<b>IVH</b>	Índice do Volume Hipocampal
<b>LASSO</b>	Least Absolute Shrinkage and Selection Operator
<b>LONI</b>	Laboratory of Neuro Imaging
<b>LSTM</b>	Long Short-Term Memory
<b>LTM</b>	Lobo Temporal Médio
<b>MAC</b>	Mapas de Ativação de Classe
<b>MC-PUAG</b>	Memory Clinic of the Policlinico Universitario “A. Gemelli”
<b>MCC</b>	Maximal Correlation Coefficient
<b>MEEM</b>	Mini-Exame do Estado Mental
<b>MHA</b>	Atenção Multi-Cabeça
<b>MIC</b>	Mutual Information Correlation
<b>MN</b>	Medidas Neuropsicológicas
<b>MRF</b>	Campo Aleatório de Markov
<b>MRF</b>	Markov Random Field
<b>mRMR</b>	minimum Redundancy Maximum Relevance
<b>MTA</b>	Atrofia do LTM
<b>N3</b>	Nonparametric Nonuniform Intensity Normalization
<b>NAC</b>	Neuroimage Analysis Center
<b>NCV</b>	Nested Cross-Validation
<b>NfL</b>	Neurofilamento leve

<b>NGTDM</b>	Matriz de Diferença de Tonalidade de Níveis de Cinza Vizinhos
<b>NLM</b>	Non-local Means
<b>PCA</b>	Análise de Componentes Principais
<b>PET</b>	Tomografia por Emissão de Pósitrons
<b>PHC</b>	Córtex Parahipocampal
<b>PMC</b>	Perceptron Multicamada
<b>PSM</b>	Preocupação Significativa com a Memória
<b>RF</b>	Radiofrequência
<b>RM</b>	Ressonância Magnética
<b>RNC</b>	Rede Neural Convolutacional
<b>ROBEX</b>	Robust Brain Extraction
<b>ROC</b>	Receiver Operating Characteristic
<b>ROI</b>	Região de Interesse
<b>RPC</b>	Razão das Curvaturas Principais
<b>SHAP</b>	SHapley Additive exPlanations
<b>SimpleITK</b>	Simple Insight Segmentation and Registration Toolkit
<b>SNR</b>	Relação sinal-ruído
<b>SVM</b>	Máquinas de Vetores de Suporte
<b>SVMcog</b>	SVM utilizando estruturas cerebrais e informações da avaliação cognitiva MEEM
<b>SVMest</b>	SVM utilizando somente estruturas cerebrais
<b>SyN</b>	Symmetric Normalization
<b>TA</b>	Temporal Anterior
<b>TAVR</b>	Teste Auditivo Verbal de Rey
<b>TE</b>	Tempo de Eco
<b>TIF</b>	Transformada Inversa de Fourier

<b>TR</b>	Tempo de Repetição
<b>VBM</b>	Morfometria Baseada em Voxels
<b>WM</b>	Substância Branca



---

# Sumário

---

<b>1</b>	<b>INTRODUÇÃO . . . . .</b>	<b>23</b>
	<i>Este capítulo apresenta o contexto no qual a pesquisa está inserida, indicando o problema a ser investigado e a motivação para resolvê-lo, além de servir como introdução ao conteúdo abordado nos próximos capítulos.</i>	
<b>1.1</b>	<b>Contexto e motivação . . . . .</b>	<b>23</b>
<b>1.2</b>	<b>Objetivos . . . . .</b>	<b>25</b>
1.2.1	Objetivo Geral . . . . .	25
1.2.2	Objetivos Específicos . . . . .	25
<b>1.3</b>	<b>Organização do trabalho . . . . .</b>	<b>26</b>
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA . . . . .</b>	<b>27</b>
	<i>Este capítulo apresenta alguns conceitos básicos complementares sobre a doença de Alzheimer, ressonância magnética e algumas técnicas de processamento de imagens para auxiliar no entendimento dos próximos capítulos deste trabalho.</i>	
<b>2.1</b>	<b>Doença de Alzheimer . . . . .</b>	<b>27</b>
2.1.1	Sintomas . . . . .	28
2.1.2	Causas . . . . .	28
2.1.3	Diagnóstico . . . . .	30
2.1.4	Tratamento . . . . .	33
<b>2.2</b>	<b>Ressonância Magnética . . . . .</b>	<b>35</b>
2.2.1	Princípios físicos . . . . .	35
2.2.2	Formação da Imagem . . . . .	39
2.2.3	Remoção do crânio . . . . .	43
2.2.4	Ruído . . . . .	44
2.2.5	<i>Bias Field</i> . . . . .	47
<b>2.3</b>	<b>Extração de atributos de textura . . . . .</b>	<b>47</b>

2.3.1	<i>Gray Level Co-occurrence Matrix (GLCM)</i> . . . . .	48
2.3.2	Redução de ruído para extração de atributos de textura . . . . .	49
2.3.3	Efeito do <i>scanner</i> . . . . .	50
<b>2.4</b>	<b>Classificação supervisionada</b> . . . . .	<b>54</b>
2.4.1	Árvore de Decisão . . . . .	54
2.4.2	Florestas Aleatórias . . . . .	57
<b>2.5</b>	<b>Considerações finais</b> . . . . .	<b>59</b>
<b>3</b>	<b>REVISÃO BIBLIOGRÁFICA</b> . . . . .	<b>61</b>
	<i>Este capítulo apresenta uma revisão das diversas técnicas propostas na literatura para a extração de características de imagens de RM baseada em engenharia de atributos visando a distinção entre CCLp e CCLe para auxiliar no diagnóstico da doença de Alzheimer.</i>	
<b>3.1</b>	<b>Trabalhos correlatos com engenharia de atributos</b> . . . . .	<b>61</b>
<b>3.2</b>	<b>Trabalhos correlatos com aprendizagem profunda</b> . . . . .	<b>71</b>
3.2.1	Discussão e Considerações Finais . . . . .	77
<b>4</b>	<b>MATERIAIS E MÉTODOS</b> . . . . .	<b>81</b>
	<i>Este capítulo apresenta uma descrição das bases de dados que foram utilizadas, bem como das técnicas que foram empregadas neste trabalho, visando facilitar a compreensão deste trabalho.</i>	
<b>4.1</b>	<b>Visão Geral</b> . . . . .	<b>81</b>
<b>4.2</b>	<b>Base de dados</b> . . . . .	<b>81</b>
4.2.1	ADNI . . . . .	81
4.2.2	NAC Brain Atlas . . . . .	87
<b>4.3</b>	<b>Pré-processamento</b> . . . . .	<b>88</b>
4.3.1	Remoção do crânio . . . . .	88
4.3.2	Redução de ruído . . . . .	89
4.3.3	Correção de <i>bias field</i> . . . . .	89
4.3.4	Padronização de histogramas . . . . .	90
4.3.5	Harmonização espacial . . . . .	90
<b>4.4</b>	<b>Segmentação de tecidos</b> . . . . .	<b>92</b>
<b>4.5</b>	<b>Parcelamento do cérebro em ROIs</b> . . . . .	<b>94</b>
<b>4.6</b>	<b>Extração de atributos</b> . . . . .	<b>95</b>
4.6.1	Volume das ROIs . . . . .	95
4.6.2	Espessura cortical . . . . .	95
4.6.3	Atributos de textura . . . . .	96
<b>4.7</b>	<b>Correção do efeito do <i>scanner</i></b> . . . . .	<b>97</b>
<b>4.8</b>	<b>Modelo de classificação</b> . . . . .	<b>97</b>
4.8.1	Seleção de atributos . . . . .	97

4.8.2	Seleção do modelo de classificação . . . . .	98
4.8.3	Treinamento do modelo . . . . .	98
4.8.4	Métricas de avaliação . . . . .	100
<b>4.9</b>	<b>Considerações finais . . . . .</b>	<b>102</b>
<b>5</b>	<b>RESULTADOS E DISCUSSÕES . . . . .</b>	<b>103</b>
	<i>Este capítulo apresenta e discute os resultados obtidos nesta pesquisa, como o pré-processamento das imagens, atributos extraídos e os experimentos de classificação das imagens de RM em CCLe e CCLp.</i>	
<b>5.1</b>	<b>Pré-processamento . . . . .</b>	<b>103</b>
5.1.1	Remoção do crânio . . . . .	103
5.1.2	Redução de ruído . . . . .	104
5.1.3	Correção de <i>bias field</i> . . . . .	104
5.1.4	Padronização de histograma . . . . .	104
5.1.5	Corregistro . . . . .	105
5.1.6	Segmentação dos tecidos cerebrais . . . . .	107
5.1.7	Segmentação das regiões de interesse . . . . .	107
<b>5.2</b>	<b>Extração de atributos . . . . .</b>	<b>108</b>
5.2.1	Volume total intracraniano . . . . .	108
5.2.2	Volume das ROIs . . . . .	110
5.2.3	Volume de cada tecido . . . . .	112
5.2.4	Imagem da espessura cortical . . . . .	114
5.2.5	Espessura cortical total . . . . .	116
5.2.6	Espessura cortical por ROI . . . . .	117
5.2.7	Atributos de textura da imagem toda . . . . .	119
5.2.8	Atributos de textura de cada ROI . . . . .	123
<b>5.3</b>	<b>Harmonização ComBat . . . . .</b>	<b>126</b>
<b>5.4</b>	<b>Experimentos de classificação . . . . .</b>	<b>127</b>
<b>5.5</b>	<b>Comparação com trabalhos da literatura . . . . .</b>	<b>133</b>
<b>5.6</b>	<b>Considerações finais . . . . .</b>	<b>135</b>
<b>6</b>	<b>CONCLUSÕES . . . . .</b>	<b>139</b>
	<i>Este capítulo sintetiza as principais conclusões do estudo do estado da arte em técnicas para a classificação de CCLp e CCLe.</i>	
<b>6.1</b>	<b>Conclusões . . . . .</b>	<b>139</b>
<b>7</b>	<b>APÊNDICE . . . . .</b>	<b>141</b>
<b>7.1</b>	<b>Apêndice 1. Atributos de textura de GLCM . . . . .</b>	<b>141</b>
<b>7.2</b>	<b>Apêndice 2. Abreviações usadas nas figuras . . . . .</b>	<b>145</b>
	<b>REFERÊNCIAS . . . . .</b>	<b>147</b>



---

# Capítulo 1

## Introdução

---

*Este capítulo apresenta o contexto no qual a pesquisa está inserida, indicando o problema a ser investigado e a motivação para resolvê-lo, além de servir como introdução ao conteúdo abordado nos próximos capítulos.*

### 1.1 Contexto e motivação

A Doença de Alzheimer (DA) é uma doença cerebral progressiva e irreversível caracterizada pela formação de neurofibrilas devido ao acúmulo anormal das proteínas beta-amiloide e tau fosforilada. Esse processo desencadeia a deterioração dos neurônios, inicialmente daqueles ligados à memória, cognição e comportamento, podendo evoluir e tornar o paciente mais vulnerável a outras doenças, que podem levá-lo a óbito (Alzheimer's Association, 2023; PORSTEINSSON et al., 2021).

Atualmente, a demência afeta cerca de 50 milhões de pessoas em todo o mundo (Alzheimer's Disease International, 2023), com a DA sendo responsável por 60-80% desses casos (Alzheimer's Association, 2023). Estima-se que até 2050 esse número possa alcançar 139 milhões, principalmente devido ao crescimento e envelhecimento da população mundial, fazendo com que os custos associados à demência aumente de US\$1,3 trilhões em 2019 para US\$2,8 trilhões até 2030 (Alzheimer's Disease International, 2023), consolidando-se como uma questão crítica de saúde pública.

Segundo Alzheimer's Association (2023), todos os pacientes diagnosticados com DA passam inicialmente pelo estágio de Comprometimento Cognitivo Leve (CCL), uma condição clínica intermediária caracterizada por sutis alterações na função cognitiva que excedem o esperado para a idade, porém não impedem a realização das atividades diárias. Alguns pacientes com CCL experimentam uma progressão gradual das dificuldades cognitivas ao longo do tempo, potencialmente evoluindo para demência, o que é denominado de CCL progressivo (CCLp). Em contrapartida, no CCL estável (CCLe), as

alterações cognitivas se mantêm relativamente constantes ao longo do tempo, sem uma evolução significativa para demência. Em ambas as variantes, os indivíduos com CCL podem enfrentar desafios relacionados à memória, linguagem, raciocínio ou outras funções cognitivas, contudo, essas dificuldades geralmente não interferem de forma substancial nas atividades cotidianas.

Apesar de não haver cura para a DA, existem tratamentos que podem retardar sua progressão e melhorar a qualidade de vida dos pacientes. Esses tratamentos incluem o uso de medicamentos, terapias cognitivas e mudanças comportamentais (Alzheimer's Association, 2023). No entanto, a eficácia desses tratamentos depende da fase da doença e das características individuais de cada paciente, sendo mais efetivos quando iniciados antes de ocorrerem danos irreversíveis no cérebro (GAO et al., 2020).

Assim, a identificação precoce de pacientes que podem desenvolver DA facilitaria o desenvolvimento e a aplicação de tratamentos e permitiria o melhor gerenciamento da doença (GAO et al., 2020). Entretanto, o diagnóstico da DA é complexo devido à sobreposição de sintomas e condições clínicas com outras doenças. Geralmente, ele envolve uma análise abrangente de biomarcadores, neuroimagens, testes neuropsicológicos, histórico familiar e informações genéticas (ATRI, 2019).

Entre as abordagens de diagnóstico, destaca-se a análise de imagens de Ressonância Magnética (RM), amplamente empregada devido à sua natureza não invasiva e à boa distinção entre tecidos moles. Isso permite a visualização de alterações sutis nos tecidos cerebrais, auxiliando no diagnóstico da DA (de MENDONÇA; FERRARI, 2023). No entanto, a identificação dessas alterações ainda apresenta um caráter subjetivo, uma vez que podem não ser perceptíveis a olho nu e estão suscetíveis à variabilidade entre operadores.

Nesse contexto, a radiômica tem se destacado em pesquisas recentes como uma abordagem inovadora que integra medicina e ciência da computação para analisar imagens médicas, capturando a heterogeneidade da doença de maneira não invasiva e econômica (TIMMEREN et al., 2020). Inicialmente muito usada na oncologia, essa técnica agora tem sido amplamente aplicada no diagnóstico e avaliação de CCL e DA (SHU et al., 2021). A radiômica quantifica características de imagem imperceptíveis ao olho humano, utilizando análises matemáticas avançadas para fornecer dados adicionais que complementam a interpretação clínica, sem substituir os métodos diagnósticos tradicionais (TIMMEREN et al., 2020).

Diante disso, pesquisas recentes têm focado no desenvolvimento de modelos baseados em engenharia de atributos (SHU et al., 2021; XIAO et al., 2021; MINHAS et al., 2021; SYAIFULLAH et al., 2021; CALANDRELLI et al., 2022; XIE et al., 2023; LEE; LEE; KIM, 2020) e aprendizagem profunda (GAO et al., 2020; BAE et al., 2021; KUNG et al., 2021; ASHTARI-MAJLAN; SEIFI; DEHSHIBI, 2022; LUO et al., 2023) capazes de reconhecer padrões nas imagens de RM, incorporando ou não outras variáveis, como pontuações clínicas, idade e sexo, a fim de classificar os indivíduos em CCLe e CCLp.

No entanto, a heterogeneidade metodológica entre os estudos e a escassez de abordagens que considerem, de forma integrada, múltiplos tipos de variáveis estruturais dificultam uma compreensão abrangente e consistente dos mecanismos associados à progressão da doença. Além disso, a quantidade limitada de imagens de pacientes com CCL disponíveis atualmente impõe restrições importantes ao desenvolvimento de modelos preditivos robustos, inviabilizando o uso de técnicas de aprendizado profundo, que requerem grandes volumes de dados para a estimação de um número elevado de parâmetros — o que pode levar ao sobreajuste. Diante desse cenário, esta pesquisa propõe o desenvolvimento de um modelo baseado em Floresta Aleatória, utilizando uma assinatura radiômica composta por atributos de espessura cortical, volume e textura extraídos de diferentes regiões cerebrais, visando prever a progressão para DA com maior reprodutibilidade.

## 1.2 Objetivos

Esta seção apresenta os objetivos geral e específicos que nortearam essa pesquisa.

### 1.2.1 Objetivo Geral

O principal objetivo dessa pesquisa foi desenvolver um método para identificar uma assinatura radiômica integrada que classificasse imagens de RM estrutural de cérebros em CCLp e CCLe, auxiliando no prognóstico ao prever se o paciente progredirá ou não para demência causada pela DA.

### 1.2.2 Objetivos Específicos

- a) Pré-processar as imagens utilizando técnicas de remoção do crânio, redução de ruído, correção de *bias field*, padronização de histogramas, orientação e correção, usando uma imagem padrão como referência (*template*);
- b) Segmentar as imagens quanto aos tipos de tecido;
- c) Estudar e identificar as regiões cerebrais de interesse e parcelar as imagens para selecionar essas regiões a partir de um atlas;
- d) Extrair medidas de volume das regiões de interesse;
- e) Extrair medidas de espessura cortical;
- f) Estudar e selecionar o(s) método(s) mais apropriados para extrair medidas de textura de imagens de RM;
- g) Estudar e selecionar o método mais apropriado para mitigar o efeito do *scanner* sobre as medidas extraídas;

- h) Estudar e selecionar o método mais apropriado para seleção de atributos e classificação.

## 1.3 Organização do trabalho

Este trabalho está dividido em seis capítulos:

- Capítulo 1: aborda o contexto e as motivações para a pesquisa.
- Capítulo 2: contém informações relacionadas à DA, imagens de RM e atributos de textura derivados de imagens de RM.
- Capítulo 3: contém a revisão bibliográfica do estado da arte relacionado à predição de conversão de casos CCL para DA.
- Capítulo 4: apresenta os materiais e métodos que foram utilizados nessa pesquisa, abrangendo a base de dados, algoritmos e técnicas.
- Capítulo 5: apresenta os resultados obtidos e discussões.
- Capítulo 6: apresenta as conclusões finais.

---

## Capítulo 2

# Fundamentação Teórica

---

*Este capítulo apresenta alguns conceitos básicos complementares sobre a doença de Alzheimer, ressonância magnética e algumas técnicas de processamento de imagens para auxiliar no entendimento dos próximos capítulos deste trabalho.*

### 2.1 Doença de Alzheimer

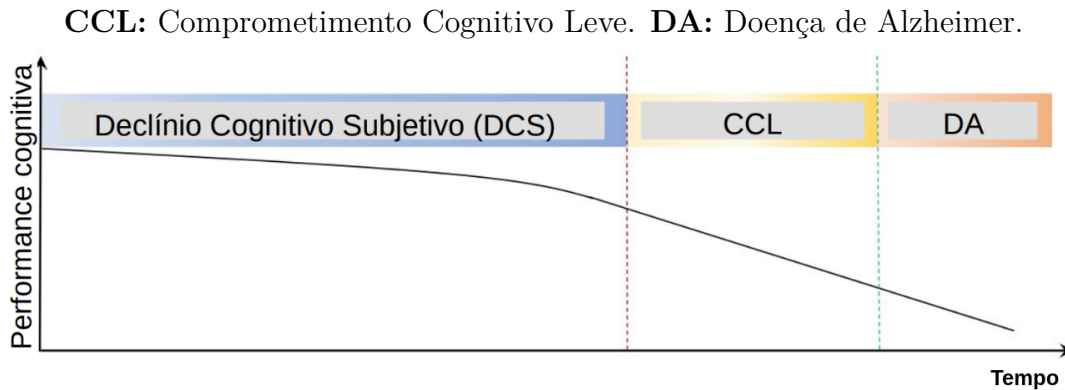
Em 1907, o psiquiatra Alois Alzheimer descreveu pela primeira vez, na Alemanha, um caso atípico de demência senil em uma mulher de 51 anos chamada Auguste Deter. A demência de Deter apresentava uma evolução progressivamente lenta, com quadros de delírios seguidos por declínio cognitivo. A partir de 1912, essa condição passou a ser chamada de Doença de Alzheimer (BRUCKI; CARAMELLI; NITRINI, 2015; POIRIER; GAUTHIER, 2016).

Conforme mostrado na Figura 1, há três estágios de declínio cognitivo: Declínio Cognitivo Subjetivo (DCS), CCL e demência (SMID et al., 2022).

O DCS refere-se à autopercepção da presença de um declínio cognitivo sem sintomas que interferem nas atividades de vida diária e com desempenho normal em testes cognitivos. Essa síndrome pode ser estável, reversiva ou progredir para um CCL, pois o DCS pode ser causado por diversos fatores como envelhecimento normal, depressão, ansiedade e uso de medicações psicoativas (SMID et al., 2022).

Por sua vez, o CCL é uma condição clínica intermediária que pode ser precedida por DCS. Para ser classificado como CCL, o paciente deve apresentar declínio cognitivo em relação ao ano anterior, superior ao esperado para sua idade, mas que não impacte nas atividades diárias (SMID et al., 2022). A partir da condição de CCL, 26% dos pacientes podem reverter o quadro; no entanto, alguns podem permanecer com o quadro clínico es-

Figura 1 – Estágios da progressão da DA.



Fonte: Extraído e adaptado de Guan et al. (2023).

tável, apresentando um CCLe, enquanto outros podem progredir para DA, experienciando um CCLp (Alzheimer's Association, 2023).

Além disso, um paciente com CCL pode ser subclassificado de acordo com o estágio em que se encontra. Pacientes com déficit cognitivo relativamente leve são classificados como Comprometimento Cognitivo Leve Inicial (CCLI), enquanto aqueles com um grau mais severo de declínio cognitivo, ainda abaixo do nível de demência, são categorizados como Comprometimento Cognitivo Leve Tardio (CCLT) (LIN et al., 2022).

### 2.1.1 Sintomas

Todos os pacientes com demência passam previamente pelo estágio de CCL, com cerca de 15% convertendo para DA depois de 2 anos e 33% dentro de 5 anos (Alzheimer's Association, 2023). Além disso, sabe-se que as alterações cerebrais começam a ocorrer aproximadamente 20 anos antes de aparecerem os primeiros sintomas e que, após o diagnóstico, a expectativa de vida média é de quatro a oito anos, podendo estender até vinte anos (Alzheimer's Association, 2023).

Na Tabela 1, é possível visualizar os principais sintomas da DA, ressaltando que esses sintomas variam de acordo com cada paciente (Alzheimer's Association, 2023).

### 2.1.2 Causas

De modo geral, a prevalência de demência em uma população está fortemente associada à idade, fatores genéticos, nível de escolaridade, fatores nutricionais e riscos de doenças vasculares cerebrais, como *diabetes mellitus*, tabagismo e hipertensão arterial. No entanto, nenhum desses fatores isoladamente é universalmente aceito como causa definitiva, configurando a etiologia da DA como multifatorial (SANVITO; ROCHA, 2008; BRUCKI; CARAMELLI; NITRINI, 2015; DINIZ; UTINO, 2023; Alzheimer's Association, 2023).

Tabela 1 – Sintomas da DA.

Estágios da DA	Sintomas
Leve	Independente nas atividades diárias, mas necessita de mais tempo para realizá-las; Dificuldades em lidar com finanças e pagamento de contas.
Moderado	Comprometimento da memória e linguagem; Confusão mental; Dificuldades em atividades realizadas por etapas, como tomar banho e se vestir; Começa a apresentar mudanças de personalidade e comportamento; Começa a ter problemas em reconhecer pessoas íntimas.
Severo	Comunicação verbal altamente comprometida; Requer cuidado 24h/dia; Movimentos comprometidos, necessitando ficar acamado; Dificuldades graves para engolir, necessitando de sondas; Devido ao uso de sondas e repouso prolongado, o paciente fica mais suscetível a inflamações e infecções.

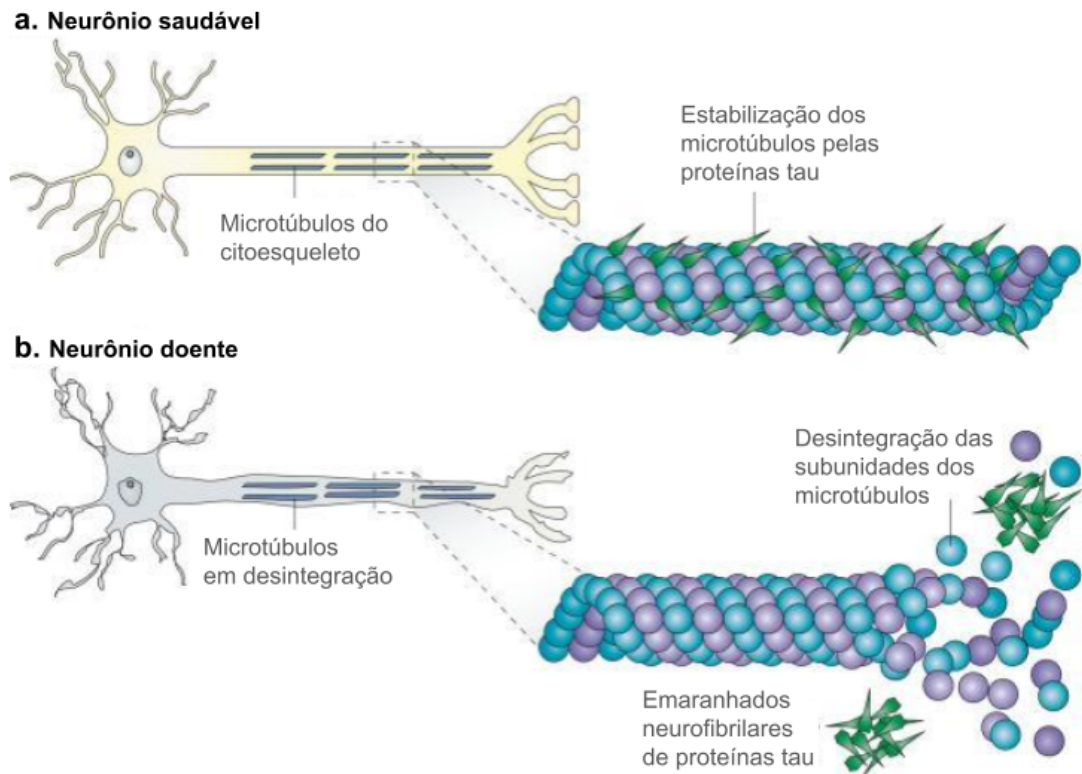
Fonte: Informações extraídas de Alzheimer's Association (2023).

Do ponto de vista microscópico, uma das principais características da DA é o depósito extracelular de peptídeos  $\beta$ -amiloides insolúveis, resultantes da clivagem da proteína amiloide (SANVITO; ROCHA, 2008; TORRES et al., 2012). Esses fragmentos peptídicos são altamente tóxicos devido à sua capacidade de se agregarem e formarem placas neuríticas, também conhecidas como placas senis. O acúmulo dessas placas pode desencadear inflamação, prejudicar a comunicação entre as células nervosas e causar perda neuronal (TORRES et al., 2012).

Outro biomarcador da DA são os emaranhados neurofibrilares, compostos pela proteína tau hiperfosforilada. A função normal da tau é estabilizar os microtúbulos que compõem o citoesqueleto das células, mantendo-os intactos (SANVITO; ROCHA, 2008; TORRES et al., 2012). No entanto, quando ocorre hiperfosforilação, a tau se agrupa com outras moléculas de tau, formando os emaranhados neurofibrilares (conforme Figura 2) (SANVITO; ROCHA, 2008; TORRES et al., 2012; BRUCKI; CARAMELLI; NITRINI, 2015; DINIZ; UTINO, 2023). Esse processo compromete a capacidade da tau de se ligar aos microtúbulos, levando a disfunções no transporte neuronal e, eventualmente, à morte das células nervosas (TORRES et al., 2012; DINIZ; UTINO, 2023).

Anatomicamente, a DA é uma doença neurodegenerativa progressiva caracterizada por um atrofia cortical difusa, acometendo principalmente a porção medial dos lobos frontais e temporais, especialmente as regiões do hipocampo, além de apresentar dilatação dos ventrículos (SANVITO; ROCHA, 2008; BRUCKI; CARAMELLI; NITRINI, 2015; DINIZ; UTINO, 2023). Mostrado na Figura 3, o encéfalo é composto por três componentes principais: a substância cinzenta (do inglês, *gray matter*, GM), formada pelos corpos

Figura 2 – Proteína tau em neurônios saudáveis e em DA.



Fonte: Imagem extraída e modificada de Brunden, Trojanowski e Lee (2009).

celulares dos neurônios presentes no córtex; a substância branca (do inglês, *white matter*, WM), constituída pelos axônios e suas bainhas de mielina, que lhe conferem a coloração branca; e o Líquido Cefalorraquidiano (do inglês, *Cerebrospinal Fluid*, CSF), um fluido que envolve o cérebro e está presente nos ventrículos (da COSTA et al., 2019).

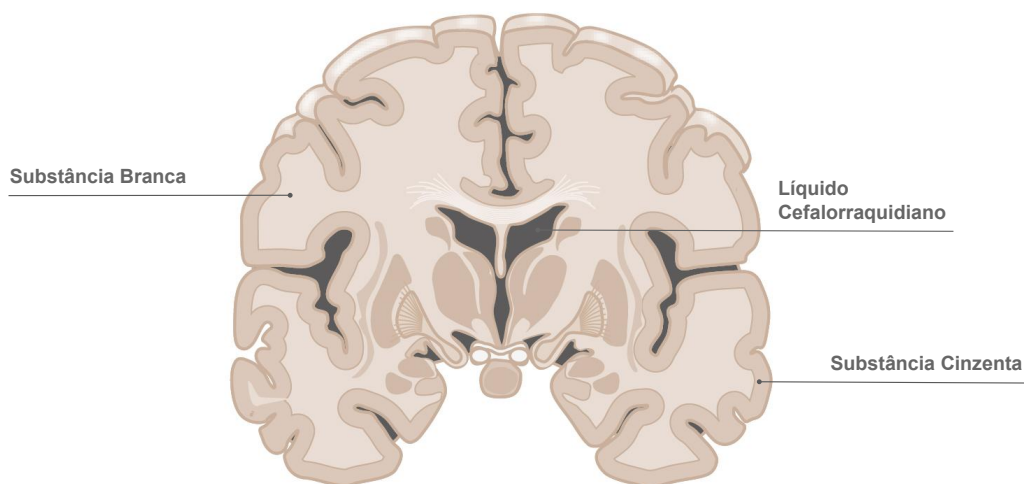
Na Figura 4, é possível observar quatro lobos do córtex cerebral e suas respectivas funções. As áreas mais afetadas pela DA são aquelas ligadas à memória, planejamento, atenção, tomada de decisões, julgamento, resposta emocional e identificação de objetos, como evidenciado na Figura 5.

De acordo com Sanvito e Rocha (2008), a DA geralmente tem início no neocórtex, a região do cérebro responsável por funções cognitivas complexas, como linguagem, pensamento abstrato, planejamento e tomada de decisões. Além disso, áreas como o hipocampo e a amígdala, associadas às funções emocionais, comportamentais e de memória, também são afetadas precocemente. À medida que a doença progride, áreas motoras e sensoriais podem ser impactadas.

### 2.1.3 Diagnóstico

O diagnóstico da DA é complexo e envolve a exclusão de outros tipos de demências com base em uma associação de exames clínicos, laboratoriais e de neuroimagem, não

Figura 3 – Corte coronal para visualização da substância cinzenta, substância branca e Líquido Cefalorraquidiano.



Fonte: Imagem extraída e modificada de NZIMBI (2024) (com licença de uso\*). \*"Slagter - Drawing Coronal section of the brain - no labels" at AnatomyTOOL.org by Ron Slagter, license: Creative Commons Attribution-NonCommercial-ShareAlike.

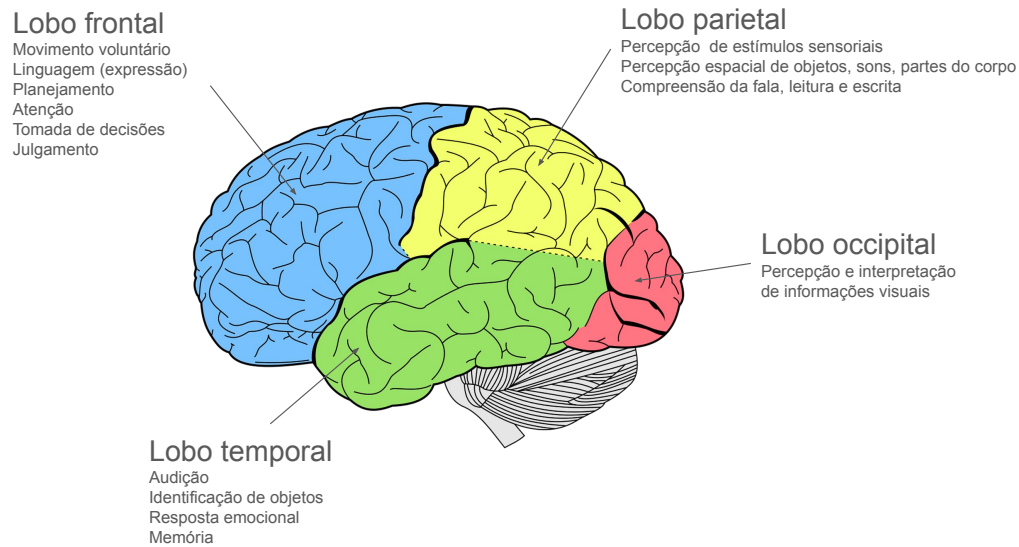
existindo um exame único que possa fornecer um diagnóstico definitivo (SANVITO; ROCHA, 2008; DINIZ; UTINO, 2023). Para diagnosticar a DA, é necessário comprovar objetivamente o comprometimento de pelo menos dois domínios cognitivos (mencionados na Tabela 2) e avaliar o impacto no desempenho das atividades diárias do paciente (BRUCKI; CARAMELLI; NITRINI, 2015; SMID et al., 2022).

Para isso, inicialmente são realizados alguns testes padronizados (como os mostrados na Tabela 3) para avaliar a presença de um possível quadro demencial no paciente, baseando-se na avaliação da cognição e do comportamento, sempre levando em consideração a idade e o nível de escolaridade do paciente.

De modo geral, esses testes de avaliações cognitivas incluem atividades que envolvam repetição, reconhecimento e memorização de figuras, palavras e números, além de execução de desenhos. Outras atividades padronizadas são usadas para avaliar memória, linguagem, comportamento, atenção, função executiva, percepções e construções visuais (DINIZ; UTINO, 2023). Além disso, para obter mais subsídios para corroborar a hipótese diagnóstica, torna-se necessária a realização de alguns exames laboratoriais e de neuroimagens.

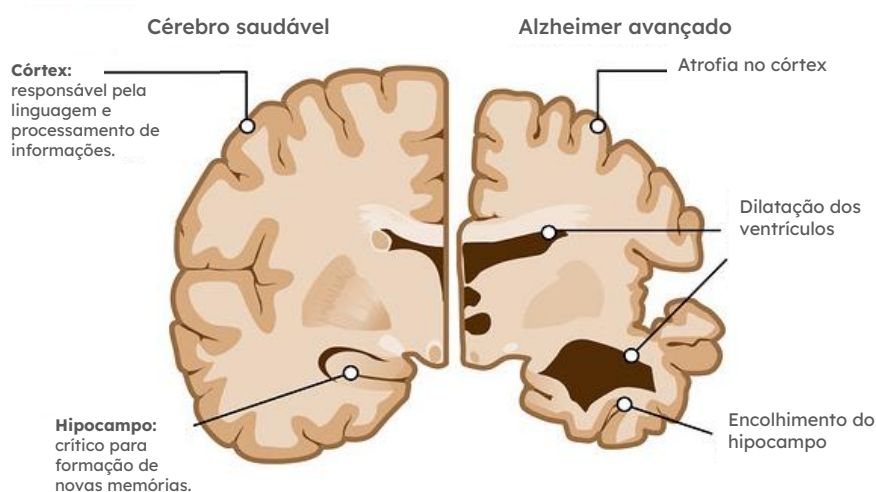
Em relação aos exames laboratoriais, inclui-se o hemograma completo, ácido fólico, vitamina  $B_{12}$ , creatinina, hormônio tireoestimulante, albumina, enzimas hepáticas, cálcio, sorologia para sífilis e, em alguns casos, sorologia para HIV (BRUCKI; CARAMELLI; NITRINI, 2015; DINIZ; UTINO, 2023).

Figura 4 – Lobos cerebrais e suas funções.



Fonte: Imagem extraída e modificada de Queensland Brain Institute (2024) com informações extraídas de Amaral e Strick (2014) e Queensland Brain Institute (2024).

Figura 5 – Comparação entre um cérebro saudável e com DA.



Fonte: Imagem extraída e modificada de Alzheimer's Association (2024).

Para a análise estrutural, são utilizados métodos de neuroimagem estrutural, como a tomografia computadorizada e a RM, sendo esta última a que permite uma avaliação mais detalhada da substância branca e da região hipocampal (BRUCKI; CARAMELLI; NITRINI, 2015). O exame de RM permite uma análise qualitativa ou quantitativa (volume) de atrofia de estruturas e de ventrículos, auxiliando na exclusão de outras doenças (SANVITO; ROCHA, 2008; DINIZ; UTINO, 2023).

Tabela 2 – Domínios cognitivos que podem ser afetados pela demência.

Domínios cognitivos	Impactos da demência
Memória	Comprometimento da capacidade para adquirir ou evocar informações recentes, com sintomas que incluem repetição das mesmas perguntas ou assuntos, esquecimento de eventos, compromissos ou do lugar onde guardou seus pertences.
Funções executivas	Comprometimento do raciocínio, da realização de tarefas complexas e do julgamento, com sintomas tais como compreensão pobre de situações de risco, redução da capacidade para cuidar das finanças, de tomar decisões e de planejar atividades complexas ou sequenciais.
Habilidades visuais-espaciais	Incapacidade de reconhecer faces ou objetos comuns, encontrar objetos no campo visual, dificuldade para manusear utensílios, para vestir-se, não explicáveis por deficiência visual ou motora.
Linguagem (expressão, compreensão, leitura e escrita)	Dificuldade para encontrar e/ou compreender palavras, erros ao falar e escrever, com trocas de palavras ou fonemas, não explicáveis por déficit sensorial ou motor.
Personalidade ou comportamento	Alterações do humor, agitação, apatia, desinteresse, isolamento social, perda de empatia, desinibição, comportamentos obsessivos, compulsivos ou socialmente inaceitáveis.

Fonte: Extraído e adaptado de Brucki, Caramelli e Nitrini (2015).

Além desses exames, há outras análises complementares, segundo Sanvito e Rocha (2008). A tomografia por emissão de pósitrons (*Positron Emission Tomography*, PET) permite evidenciar uma redução da atividade metabólica nas regiões temporoparietais de ambos os hemisférios cerebrais. A cintilografia cerebral tomográfica, por sua vez, fornece imagens que refletem a atividade metabólica e o fluxo sanguíneo cerebral, sendo capaz de demonstrar hiperfusão nas mesmas regiões temporoparietais bilateralmente. O eletroencefalograma fornece informações sobre a atividade elétrica do cérebro. Já a análise do CSF envolve a coleta e exame laboratorial de suas propriedades físico-químicas, podendo evidenciar aumento da concentração de proteína tau, alterações nos níveis da proteína amiloide e redução dos peptídeos  $\beta$ -amiloides, marcadores frequentemente associados a processos neurodegenerativos.

### 2.1.4 Tratamento

Após ser diagnosticado com DA, o paciente pode iniciar tratamentos que desaceleram a progressão da doença e melhoram sua qualidade de vida, uma vez que a DA não possui cura (DINIZ; UTINO, 2023). De modo geral, existem alguns tipos de tratamentos: farmacológicos, comportamentais e cognitivos.

A indicação de tratamento farmacológico depende do estágio da DA e da presença e

Tabela 3 – Exames de avaliação do comprometimento cognitivo.

Exame	Pontuação	Áreas de avaliação
Mini-Exame do Estado Mental (MEEM)	0 - 30 (maior pontuação indica melhor desempenho)	Orientação temporoespacial, memória, cálculo, linguagem e habilidades visuais construtivas (BRUCKI; CARAMELLI; NITRINI, 2015).
<i>Mattis Dementia Rating Scale</i> (MDRS)	0 - 144 (maior pontuação indica melhor desempenho)	Atenção, iniciação, perseveração, construção, conceituação e memória (FOSS; VALE; SPECIALI, 2005).
<i>Montreal Cognitive Assessment</i> (MoCA)	0 - 30 (maior pontuação indica melhor desempenho)	Memória, atenção, linguagem, orientação temporoespacial, habilidades visuoespaciais e executivas (Ministério da Saúde, 2024).
<i>Consortium Establish of Alzheimer Disease</i> (CERAD)	Depende do teste (maior pontuação indica melhor desempenho)	Memória, linguagem, praxia e função executiva (AZAMBUJA, 2007).
<i>Alzheimer's Disease Assessment Scale - Cognitive Subscale</i> (ADAS-Cog)	0 - 70 (menor pontuação indica melhor desempenho)	Memória, atenção, linguagem, praxia e outras (SAJJADI; BROWN, 2015).
<i>Clinical Dementia Rating</i> (CDR)	0 - 3 Sendo: 0: Saudável 0.5: Demência questionável 1: Demência leve 2: Demência moderada 3: Demência grave	Memória, orientação, juízo, resolução de problemas, atividades domésticas, <i>hobbies</i> e cuidado pessoal (AZAMBUJA, 2007).

Fonte: Autoria própria.

predisposição a outras doenças (DINIZ; UTINO, 2023; Alzheimer's Association, 2023). Os fármacos, empregados isoladamente ou em combinação, têm como objetivo regular neurotransmissores, controlar sintomas neuropsiquiátricos (podendo incluir antidepressivos, ansiolíticos e outros fármacos para controle de distúrbios comportamentais) ou retardar a progressão da doença, podendo apresentar efeitos colaterais (SANVITO; ROCHA, 2008; Alzheimer's Association, 2023). Outros tratamentos não farmacológicos podem ser aplicados para manter ou melhorar a função cognitiva do paciente e ajudar na regulação de sintomas comportamentais e psicológicos (depressão, agitação, agressividade, etc). Esses tratamentos consistem na realização de exercícios físicos e mentais, podendo envolver fonoaudiologia, psicoterapias, fisioterapias e terapias alternativas (DINIZ; UTINO, 2023; Alzheimer's Association, 2023).

Diante da ausência de cura para DA, é importante conscientizar a população sobre os

fatores de riscos para demência que podem ser reduzidos com a mudança de conduta individual, como atividade física, educação, convívio social, mentalidade ativa e o tabagismo (Alzheimer's Association, 2023).

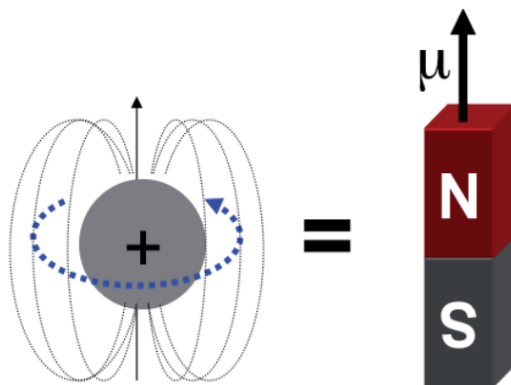
## 2.2 Ressonância Magnética

Como mencionado anteriormente, a RM é um método de diagnóstico por imagem que permite extrair informações estruturais do cérebro com detalhes anatômicos tridimensionais (3D), auxiliando na identificação da DA. A seguir, serão abordados, de forma introdutória e simplificada, alguns princípios básicos dessa técnica de diagnóstico, incluindo a obtenção e as características das imagens de RM.

### 2.2.1 Princípios físicos

Os átomos de hidrogênio possuem o núcleo mais simples, sendo composto por um único próton (uma partícula carregada positivamente), que gira em torno de seu próprio eixo, exibindo a propriedade de momento angular (ou *spin*) (MAZZOLA, 2009; HAGE; IWASAKI, 2009). Esse movimento gera um campo magnético ao redor do núcleo, fazendo-o se comportar de maneira análoga a um ímã (dipolo magnético). Este fenômeno intrínseco está associado a um momento magnético, conforme ilustra a Figura 6 (MAZZOLA, 2009; HAGE; IWASAKI, 2009).

Figura 6 – *Spin* e momento magnético ( $\mu$ ) do próton de hidrogênio.



Fonte: Imagem extraída e modificada de Mazzola (2009).

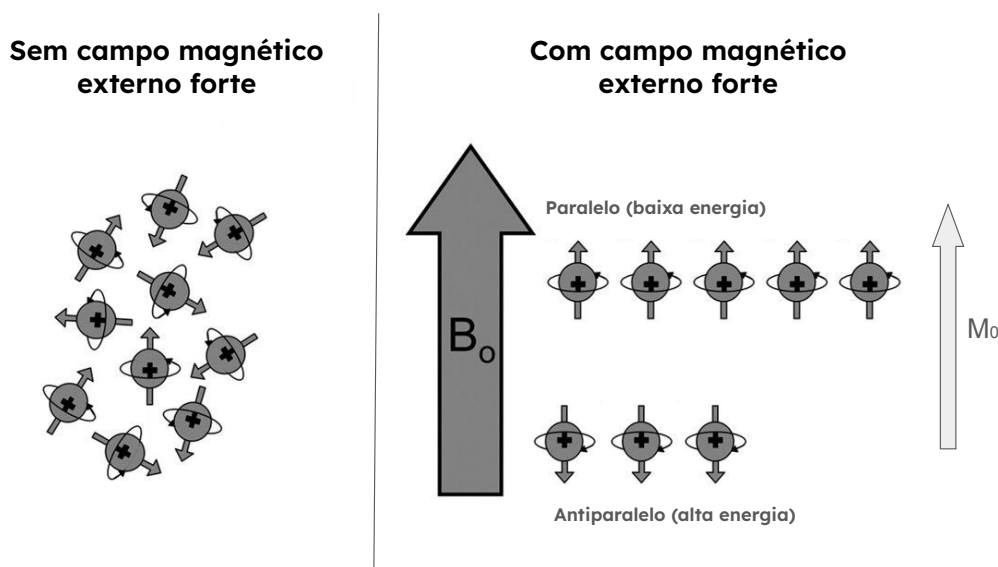
O tecido humano é composto em grande parte por átomos de hidrogênio, representando aproximadamente 10% do peso corporal (MAZZOLA, 2009). Além disso, o hidrogênio possui um momento magnético maior em relação aos outros elementos presentes nos tecidos humanos, o que lhe confere uma maior capacidade de interagir com campos magnéticos externos e, conseqüentemente, produzir um sinal de RM mais intenso (HAGE;

IWASAKI, 2009). Por isso, os átomos de hidrogênio são os mais apropriados para a obtenção de imagens por RM.

Naturalmente, os núcleos de hidrogênio apresentam orientações aleatórias no corpo humano, apresentando momento magnético resultante nulo (HAGE; IWASAKI, 2009). Porém, conforme Figura 7, ao serem expostos a um campo magnético externo forte ( $B_0$ ) (como 1,5T ou 3T), os vetores de magnetização de cada núcleo se orientam de acordo com a direção do campo aplicado, resultando em um aumento significativo do momento magnético (HAGE; IWASAKI, 2009).

Alguns *spins* se alinham paralelamente ao campo magnético externo, enquanto outros se alinham antiparalelamente, ocupando níveis de baixa e alta energia, respectivamente. Dado que a maioria dos *spins* se alinha paralelamente, o momento magnético resultante é determinado pelos momentos magnéticos dos *spins* paralelos excedentes (MAZZOLA, 2009).

Figura 7 – Alinhamento dos prótons de hidrogênio em um campo magnético externo forte e o momento magnético resultante ( $M_0$ ).

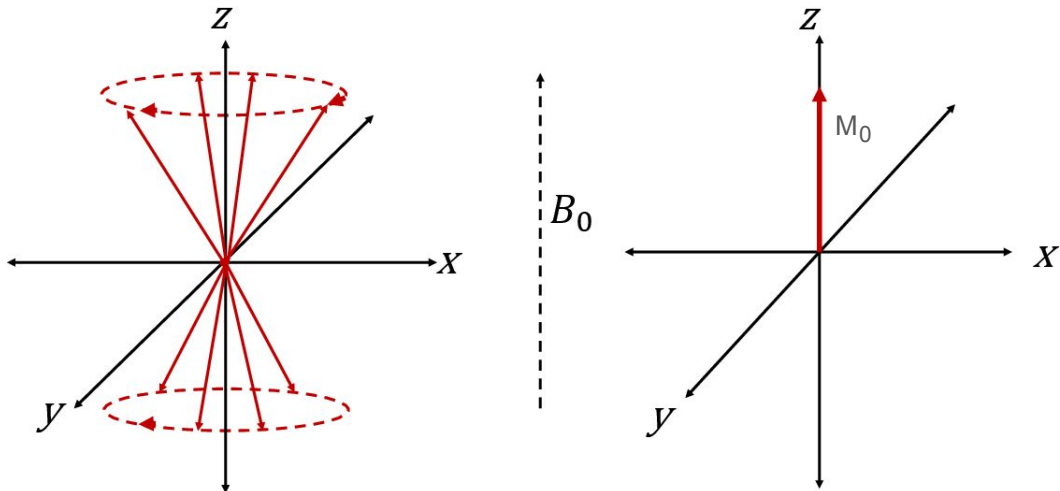


Fonte: Imagem extraída e modificada de Pooley (2005).

Como parte de sua interação com o campo magnético externo, os prótons iniciam um movimento adicional conhecido como precessão (MAZZOLA, 2009; HAGE; IWASAKI, 2009). Esse fenômeno gradualmente modifica a orientação do eixo de rotação do próton, resultando em um movimento cônico ao redor do eixo de rotação original (conforme Figura 8). A precessão ocorre em uma frequência conhecida por frequência de Larmor, que varia de acordo com a intensidade do campo magnético aplicado (MAZZOLA, 2009; HAGE; IWASAKI, 2009).

Após serem submetidos ao campo  $B_0$ , uma corrente elétrica alternada é aplicada às bobinas de radiofrequência (RF), criando um novo campo magnético temporário ( $B_1$ )

Figura 8 – Movimento de precessão do próton de hidrogênio.



Fonte: Imagem extraída e modificada de Issa (2018).

oscilante na frequência de Larmor (MAZZOLA, 2009; BUSHBERG et al., 2012) e perpendicular a  $B_0$ , perturbando os *spins*. Diante disso, todos os momentos magnéticos individuais dos *spins* são desviados em  $90^\circ$ , passando a girar em sincronia no plano transversal (plano  $xy$ ) (MAZZOLA, 2009).

Com isso, observa-se o aumento da magnetização no plano  $xy$  ( $M_{xy}$ ) e a diminuição da magnetização longitudinal ( $M_L$ ) (MAZZOLA, 2009). Isso faz com que os *spins* sincronizados criem um campo magnético oscilante perpendicular a  $B_0$ , induzindo uma corrente elétrica em uma bobina receptora dentro do *scanner* de RM, conforme a Lei de Faraday (BUSHBERG et al., 2012).

Ao cessar a emissão de pulso RF, os *spins* gradualmente liberam a energia absorvida para os *spins* vizinhos e para a rede de *spins* circundante, iniciando o processo de relaxação, retornando ao estado de equilíbrio. Com isso, o vetor de magnetização é realinhado paralelamente a  $B_0$  (MAZZOLA, 2009).

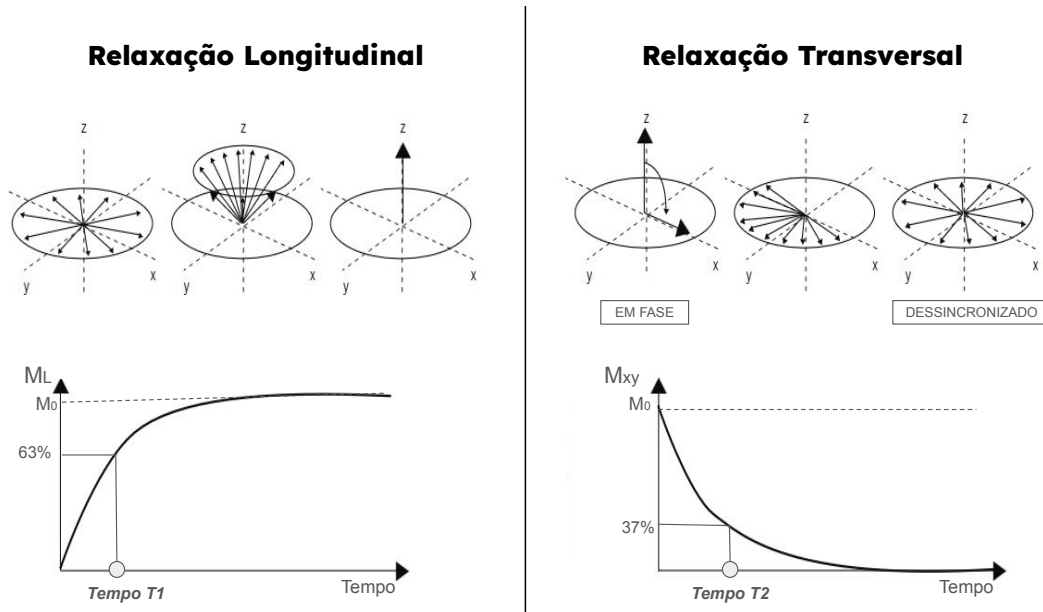
### 2.2.1.1 Ponderação das Imagens de RM

O processo de relaxação ocorre simultaneamente em duas modalidades distintas: relaxação longitudinal e relaxação transversal. Esses processos resultam em duas constantes fundamentais conhecidas como tempo T1 e tempo T2, medidas em milissegundos (MAZZOLA, 2009; HAGE; IWASAKI, 2009).

Como mostrado na Figura 9, o tempo T1 refere-se ao tempo necessário para a magnetização recuperar 63% do seu valor inicial ( $M_0$ ) em relação ao plano longitudinal (eixo  $z$ ). Já o tempo T2 representa o tempo necessário para que a magnetização alcance 37%

do seu valor inicial em relação ao plano transversal (MAZZOLA, 2009).

Figura 9 – Relaxação longitudinal e transversal.



Fonte: Imagem extraída e modificada de Cagnie et al. (2011).

Entretanto, devido à rápida dessincronização dos *spins* causada pelas interações com o ambiente circundante, o sinal decai rapidamente durante a relaxação, prejudicando a qualidade da imagem. Assim, para contornar esse problema, as sequências de pulso *spin* eco são comumente utilizadas.

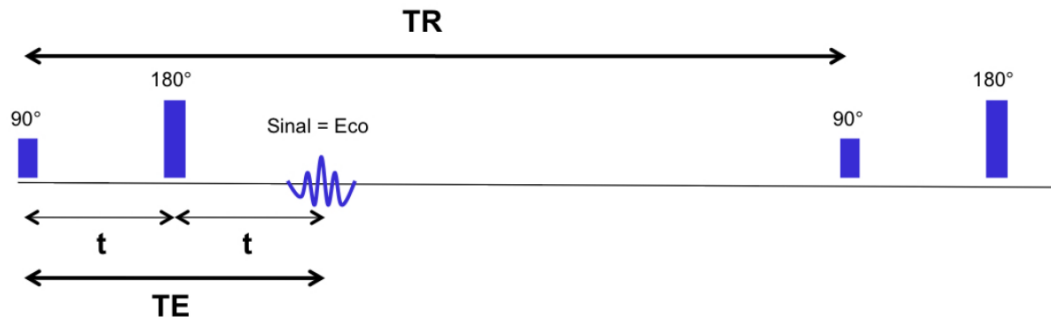
Conforme mostrado na Figura 10, a sequência de pulso *spin* eco começa com a aplicação de um pulso de RF de  $90^\circ$ . Após um certo período, aplica-se um pulso de RF de  $180^\circ$ , invertendo as orientações de magnetização dos *spins* e realinhando aqueles que estavam rapidamente entrando em dessincronização (HAGE; IWASAKI, 2009). Com isso, após um determinado período, os sensores da RM são ativados obtendo-se um sinal de eco, que reflete o estado de coerência dos *spins* (MAZZOLA, 2009; HAGE; IWASAKI, 2009).

O período de tempo entre a aplicação do pulso de RF de  $90^\circ$  e a obtenção do sinal eco é chamado de Tempo de Eco (TE), que determina o quanto de relaxação em relação ao plano longitudinal estará presente no sinal capturado pelos sensores (MAZZOLA, 2009). Essa sequência pode ser realizada repetidamente e o intervalo entre os dois pulsos iniciais de cada sequência é chamada Tempo de Repetição (TR) (MAZZOLA, 2009).

Devido às diferenças nas composições dos tecidos, como teor de água e de macromoléculas, observa-se uma diferença entre os tempos T1 e T2 do CSF, GM e WM, sendo possível obter imagens com alto contraste entre esses tecidos. Para isso, torna-se necessária a manipulação do TR e do TE para obter imagem ponderada em T1, T2 ou densidade de prótons (DP), conforme Tabela 4, sendo a ponderação em T1 a mais usada para diagnóstico de DA devido ao maior contraste anatômico entre GM e WM.

Figura 10 – Sequência de pulso *spin eco*.

**TR:** Tempo de Repetição. **TE:** Tempo de Eco.



Fonte: Imagem extraída de Mazzola (2009).

Tabela 4 – Diferentes ponderações para imagens de RM.

Tempo de repetição (TR)	Tempo de eco (TE)	Ponderação
TR Curto (< 500 ms)	TE Curto (5 a 25 ms)	T1
TR Longo (> 1500 ms)	TE Longo (> 90 ms)	T2
TR Longo (> 1500 ms)	TE Curto (5 a 25 ms)	DP

Fonte: Tabela extraída de Mazzola (2009).

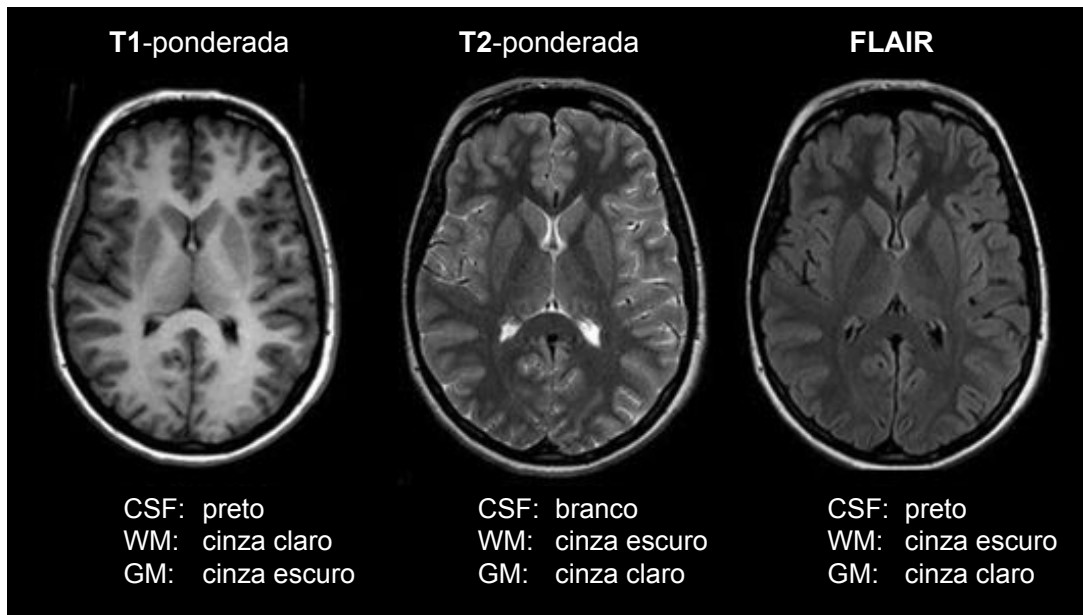
Outras técnicas complementares podem ser usadas para melhorar a qualidade das imagens, tal como a sequência *fluid-attenuated inversion recovery* (FLAIR) e uso de contraste. As imagens em FLAIR, são imagens ponderadas em T2 com o sinal do CSF suprimido, permitindo uma melhor visualização do GM em locais com contaminação de CSF, sendo muito utilizada para identificação de lesões cerebrais (ROOZPEYKAR et al., 2022). Na Figura 11, é possível observar as diferenças entre as imagens ponderadas em T1, T2 e FLAIR (HAGE; IWASAKI, 2009).

Já para obter imagens com uso de contraste, é aplicada ao paciente uma injeção intravenosa de um agente de contraste paramagnético, geralmente o gadolínio. O contraste tem a capacidade de alterar as propriedades magnéticas dos tecidos circundantes, reduzindo o tempo de relaxação T1 e, conseqüentemente, resultando em um maior sinal nessa região (ELIAS JUNIOR et al., 2008). Com isso, é capaz de realçar pequenas lesões em vasos sanguíneos, inflamações, entre outras anormalidades (ROOZPEYKAR et al., 2022).

## 2.2.2 Formação da Imagem

Uma imagem de RM é uma imagem 3D formada por imagens bidimensionais (2D) empilhadas, cada uma sendo uma fatia (ou corte). Para formar cada uma das fatias, o equipamento deve ser capaz de distinguir espacialmente os sinais recebidos. Para isso, o *scanner* do equipamento de RM contém bobinas de gradiente, que causam variações

Figura 11 – Comparação entre imagens ponderadas em T1, T2 e FLAIR.



Fonte: Imagem extraída e modificada de West et al. (2017).

lineares no campo magnético, fazendo com que a localização espacial ocorra de acordo com dois princípios: seleção de corte e preenchimento do espaço K (HANSON; GROTH, 2009).

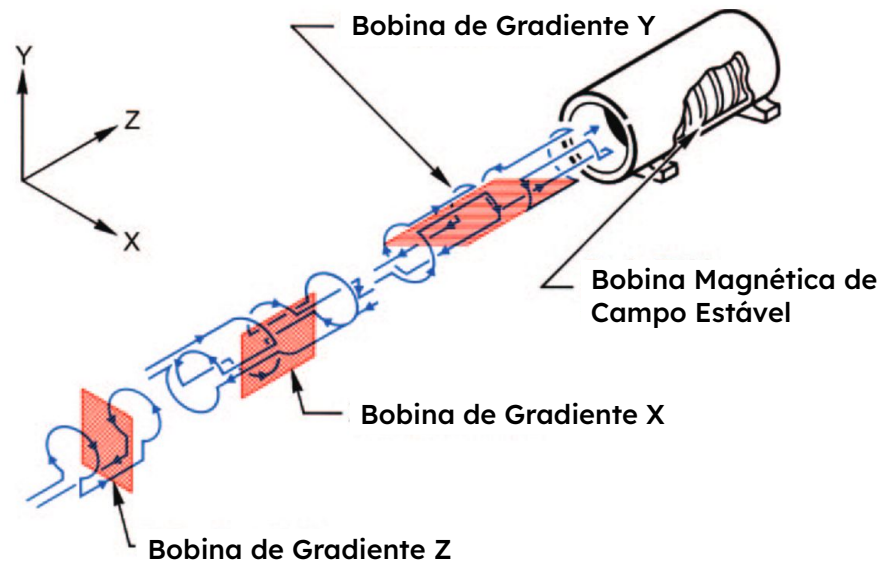
### 2.2.2.1 Seleção de corte

Ao ativar um gradiente, torna-se possível controlar a variação linear do campo magnético, o que é crucial para possibilitar uma localização espacial precisa nas três direções ( $x$ ,  $y$  e  $z$ ) (JACOBS; IBRAHIM; OUWERKERK, 2007). Conforme mencionado anteriormente, os *spins* executam o movimento de precessão ao redor de  $B_0$  com uma frequência que é determinada pelo campo magnético local. Como resultado da ativação do gradiente, cada *spin* passa a estar submetido a um campo magnético local diferente e, consequentemente, o movimento de precessão ocorre com frequência e fase distintas, refletindo a posição do núcleo dentro do equipamento de RM (JACOBS; IBRAHIM; OUWERKERK, 2007; HANSON; GROTH, 2009).

Para isso, bobinas de gradiente encontram-se posicionadas concentricamente na máquina de RM, conforme mostra a Figura 12. Essas bobinas são usadas sequencialmente para realizar três funções: seleção de corte, codificação de fase e codificação de frequência, cada uma representando o acionamento do gradiente em uma dada direção (JACOBS; IBRAHIM; OUWERKERK, 2007; MAZZOLA, 2009).

Para realizar a seleção de corte, aplica-se um gradiente de campo magnético ao longo do eixo  $z$ , fazendo com que o campo magnético varie ao longo desse eixo. Isso faz com que os *spins* comecem o movimento de precessão com frequências diferentes ao longo do

Figura 12 – Configuração das bobinas de gradiente.



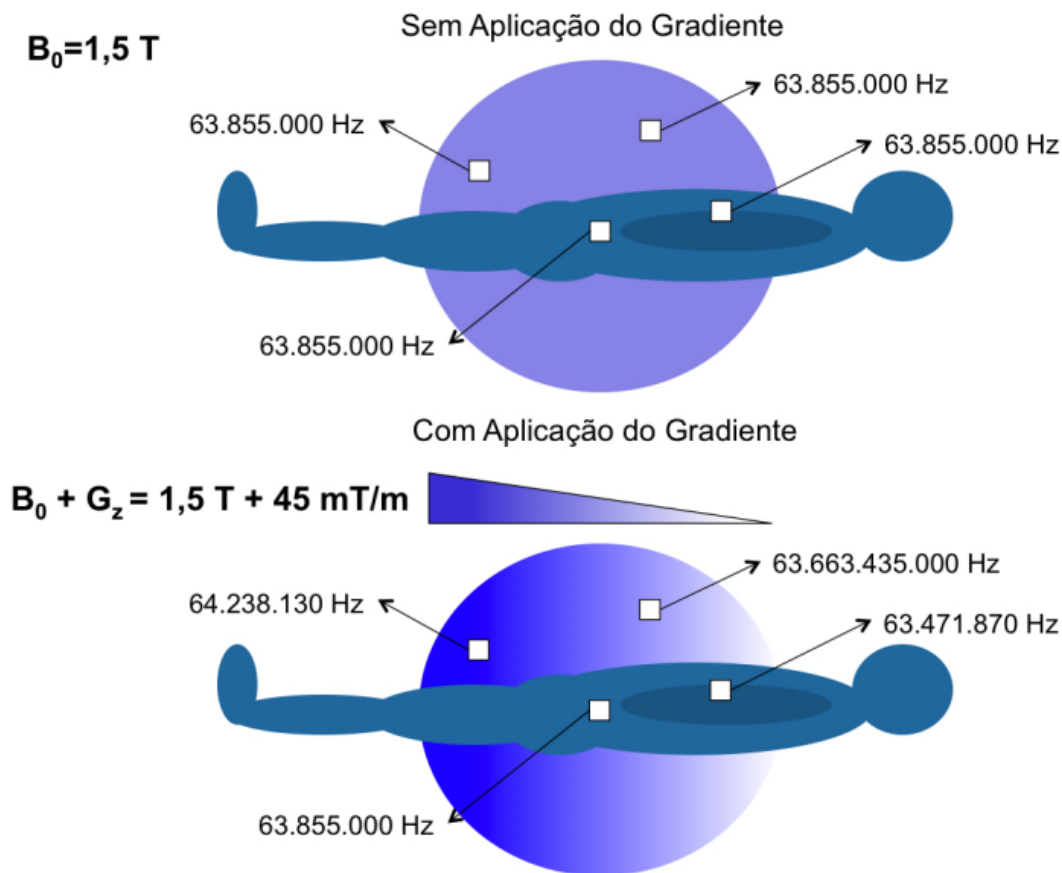
Fonte: Imagem extraída e modificada de Jacobs, Ibrahim e Ouwerkerk (2007).

eixo  $z$ , conforme representado na Figura 13 (JACOBS; IBRAHIM; OUWERKERK, 2007; MAZZOLA, 2009). Em seguida, seleciona-se a fatia específica do corpo do paciente a partir da emissão de pulso de RF na mesma frequência de precessão dos *spins* da fatia de interesse, perturbando somente esses *spins*. A espessura das fatias depende da largura de banda do pulso de RF e da intensidade do gradiente aplicado (JACOBS; IBRAHIM; OUWERKERK, 2007).

Posteriormente, aplica-se o gradiente de codificação de fase, que altera o campo magnético ao longo do eixo  $y$ . Como resultado, as fases dos *spins* da fatia selecionada são alteradas de acordo com o campo magnético local de cada um, gerando uma variação das fases dos *spins* de acordo com suas posições em relação ao eixo  $y$  (JACOBS; IBRAHIM; OUWERKERK, 2007). Dessa forma, cada *spin* gera um sinal diferente, sendo possível codificar a posição de cada um em relação ao eixo  $y$  (JACOBS; IBRAHIM; OUWERKERK, 2007; MAZZOLA, 2009). A força desse gradiente de campo magnético é ajustada para diferentes etapas da sequência de pulso (MAZZOLA, 2009).

Por fim, é necessário identificar a localização de cada núcleo dentro da fatia em relação ao eixo  $x$  para, posteriormente, formar uma imagem 3D. Para isso, realiza-se a codificação de frequência, que consiste em ativar um gradiente magnético ao longo do eixo  $x$ , alterando linearmente o campo magnético local de cada um dos *spins* ao longo desse eixo. Como resultado, os *spins* da fatia selecionada passam a realizar a precessão com frequências distintas, sendo possível codificar a posição espacial de cada um de acordo com a frequência do sinal que emitirem (JACOBS; IBRAHIM; OUWERKERK, 2007; MAZZOLA, 2009).

Figura 13 – Efeito de aplicação de um gradiente de campo magnético na direção do eixo  $z$  com amplitude 45 mT/m.



Fonte: Imagem extraída de Mazzola (2009).

### 2.2.2.2 Espaço K

Todas as informações obtidas pelos gradientes sobre uma fatia, como localização, frequência e fase são armazenadas em formato de matriz, constituindo o espaço K, permitindo, assim, a amostragem do sinal em um número reduzido de pontos em comparação com a aquisição direta no domínio do espaço real (JACOBS; IBRAHIM; OUWERKERK, 2007).

Cada linha do espaço K é preenchida com um eco (sinal) codificado a cada variação na amplitude do gradiente de codificação de fase durante uma sequência de pulso, enquanto cada ponto corresponde à amplitude do sinal recebido pela bobina em um certo instante (MAZZOLA, 2009). As coordenadas dessa matriz são o gradiente de codificação de frequência ( $k_x$ ) e o gradiente de codificação de fase ( $k_y$ ) (MAZZOLA, 2009; JACOBS; IBRAHIM; OUWERKERK, 2007).

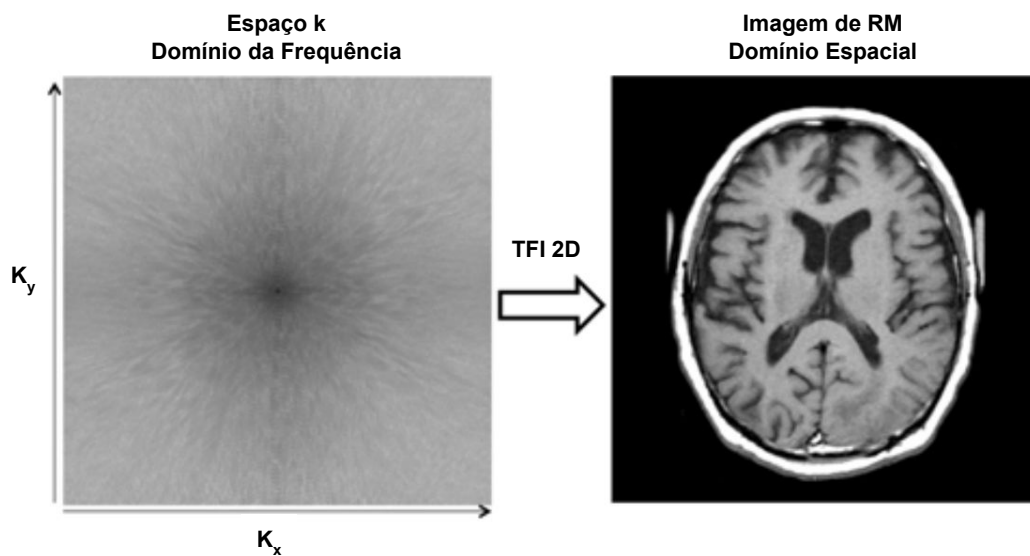
Quanto maior o número de linhas de um espaço K ou maior a quantidade de espaço K para uma imagem, maior será a quantidade de sinal coletado. Porém, o tempo necessário para adquirir uma imagem aumentará consideravelmente (MAZZOLA, 2009).

Além disso, devido à redução do sinal, causada pela ativação do gradiente, os extre-

mos do espaço  $K$  são preenchidos com sinal de baixa amplitude e estão relacionados à resolução espacial, enquanto as linhas centrais são preenchidas por sinal de maior amplitude, resultando em contraste na imagem de RM (MAZZOLA, 2009; HANSON; GROTH, 2009).

Após a coleta e o armazenamento de 64, 128, 256 ou mais ecos no espaço  $K$  (mapeados em fase e frequência), a imagem de RM pode ser construída aplicando-se a Transformada Inversa de Fourier (TIF), que passará as informações do domínio da frequência (espaço  $K$ ) para o domínio espacial (imagem), conforme Figura 14 (JACOBS; IBRAHIM; OUWERKERK, 2007; MAZZOLA, 2009; HANSON; GROTH, 2009).

Figura 14 – Espaço  $K$  e a imagem de RM correspondente após a aplicação da TIF bidimensional.



Fonte: Imagem adaptada de Mazzola (2009).

### 2.2.3 Remoção do crânio

A imagem de RM contém vários tecidos não-cerebrais, como pele, pescoço, gordura, globo ocular e músculo (PEI et al., 2022). A presença desses tecidos torna desafiadora a análise automática de imagens cerebrais. Portanto, um processamento preliminar é necessário para remover tecidos não cerebrais, chamado de remoção do crânio (*skull stripping*), o que remove o crânio e os outros tecidos não-cerebrais das imagens de RM, levando a uma melhor segmentação de diferentes regiões cerebrais e permitindo obter maior reprodutibilidade em grande escala e resultados mais precisos (REHMAN; HWANG; LEE, 2020; PEI et al., 2022).

É possível realizar a remoção do crânio manualmente, porém, depende muito tempo e é suscetível a variações devido à atividade humana, não sendo viável para um estudo

de grande escala, como o aqui proposto. Diante disso, pode-se utilizar o método *Robust Brain Extraction* (ROBEX), que possui uma abordagem híbrida, combinando modelos discriminativos e generativos (IGLESIAS et al., 2011).

O modelo discriminativo, baseado em Floresta Aleatória (*Random Forest*), é responsável por classificar cada *voxel* como parte do cérebro ou não, utilizando características extraídas da imagem, como informações de intensidade, gradientes, textura e características morfológicas dos *voxels* (IGLESIAS et al., 2011).

Além disso, o método incorpora o corregristo com atlas para alinhar as imagens com um modelo anatômico de referência, garantindo que as características anatômicas das imagens de entrada estejam consistentemente alinhadas com um modelo de referência. Durante o corregristo, a imagem de entrada é transformada geometricamente para se ajustar ao modelo de referência, utilizando técnicas de registro espacial (IGLESIAS et al., 2011).

O modelo generativo, conhecido como Modelo de Distribuição de Ponto, captura a variabilidade das formas cerebrais com base em pontos de referência identificados nas imagens. Esse modelo é construído a partir desses pontos de referência e utiliza técnicas estatísticas, como a Análise de Componentes Principais (*Principal Component Analysis*, PCA), para modelar as variações nas formas do cérebro (IGLESIAS et al., 2011). Isso permite gerar formas de cérebro realistas e adaptáveis, fornecendo uma base sólida para a segmentação.

O método também incorpora a morfologia matemática para manipular características da imagem, como bordas, regiões de interesse e ruído, utilizando operações como erosão, dilatação, abertura e fechamento para melhorar a qualidade da imagem e prepará-la para a segmentação (IGLESIAS et al., 2011). Essas operações ajudam a remover artefatos indesejados, preencher pequenos espaços vazios e realçar estruturas relevantes, contribuindo para uma segmentação mais precisa e confiável.

Por fim, o modelo deformável é utilizado para orientar a segmentação adaptável às variações anatômicas entre diferentes imagens, ajustando o modelo de segmentação para se adaptar às características específicas de cada imagem, garantindo uma segmentação precisa e robusta, mesmo em casos de variações na forma e na anatomia do cérebro (IGLESIAS et al., 2011). O método foi construído para ser invariante à intensidade e para preservar o cerebelo e o tronco encefálico, além da massa cerebral, e não requer ajuste de parâmetros.

## 2.2.4 Ruído

O principal desafio na obtenção de imagens de RM consiste nos ruídos, que podem ter origem variada, seja por influência do equipamento, fatores biológicos ou até mesmo inerentes à própria técnica de imageamento, impondo um limite superior à qualidade da

imagem, o que prejudica tanto sua precisão quanto a interpretação (HANSON; GROTH, 2009).

A relação sinal-ruído (*Signal-to-Noise ratio*, SNR) de uma imagem de RM é uma medida que descreve a proporção entre o sinal desejado, que corresponde às informações anatômicas que queremos capturar na imagem, e o ruído presente na imagem, que consiste em variações aleatórias e indesejadas que podem distorcer essas informações.

Os ruídos biológicos podem ser causados pelo movimento, respiração ou pulsação do paciente durante o exame de RM (HANSON; GROTH, 2009). Mesmo que o paciente permaneça completamente imóvel, os ruídos ainda podem surgir devido ao movimento aleatório de partículas carregadas (íons) presentes em seu corpo, algo inerente à natureza térmica do corpo humano (HANSON; GROTH, 2009). Devido à sua energia térmica, esses íons se movem de forma randômica, o que pode gerar alterações no campo magnético local, resultando na emissão de ondas de rádio aleatórias, as quais podem interferir no sinal de RM, adicionando ruído às imagens (HANSON; GROTH, 2009; AJA-FERNÁNDEZ; VEGAS-SÁNCHEZ-FERRERO, 2016).

Já em relação ao equipamento, este pode ser uma fonte significativa de ruído eletrônico, especialmente de ruído térmico. Durante a recepção dos sinais pelo *scanner* pode haver interferências elétricas, sendo que tais interferências podem ser provenientes tanto do próprio equipamento, especialmente das bobinas de RF, cujos elétrons livres apresentam movimento estocástico, quanto de fontes externas, como outros dispositivos eletrônicos próximos, cabos de alimentação ou até mesmo a própria rede elétrica (BUSHBERG et al., 2012; AJA-FERNÁNDEZ; VEGAS-SÁNCHEZ-FERRERO, 2016).

O ruído, por sua natureza, não é codificado espacialmente e, por ser recebido de forma contínua durante a aquisição da imagem, é distribuído uniformemente no espaço  $K$ , abrangendo todas as frequências. Isso resulta em uma imagem de RM com ruído distribuído uniformemente (HANSON; GROTH, 2009). A modelagem do ruído da imagem de RM é frequentemente descrita pela distribuição de probabilidade Rician devido à combinação de componentes determinísticos (sinal) e estocásticos (ruído) no processo de formação da imagem. No entanto, para valores elevados da SNR, a distribuição de probabilidade Rician pode ser aproximada pela distribuição Gaussiana. (BUSHBERG et al., 2012; AJA-FERNÁNDEZ; VEGAS-SÁNCHEZ-FERRERO, 2016).

A SNR é influenciada por vários fatores, incluindo o número de excitações da fatia, as dimensões do *voxel*, o número de amostras no espaço  $K$  e a largura de banda de recepção, assim como o tempo de aquisição da imagem (TOFTS, 2003). A SNR é diretamente proporcional à raiz quadrada do tempo de aquisição da imagem e ao tamanho do *voxel*, o qual é influenciado pela espessura da fatia e do campo de visão (região do paciente que está sendo visualizada e capturada pela máquina durante o processo de aquisição da imagem) (TOFTS, 2003).

### 2.2.4.1 Redução de ruído

Para a atenuação de ruído (*denoising*), pode ser utilizado o método *Non-local Means* (NLM), o qual itera sobre a imagem buscando por *patches* (pequenas regiões na imagem) similares para reconstruir a imagem original sem ruído, o que permite a preservação de detalhes importantes e evita a criação de artefatos (BUADES; COLL; MOREL, 2005a).

De maneira geral, o método consiste em identificar *patches* de comparação que sejam semelhantes ao *patch* de referência (considerado para remoção de ruído). Para isso, calcula-se a similaridade entre *patches* em uma região da imagem. Esse processo envolve calcular e atribuir pesos aos *pixels* da vizinhança, com base na distância entre os *patches* e na distância dos *pixels* de cada *patch* ao *pixel* central do *patch* a qual pertence, além da distância euclidiana entre as intensidades dos *pixels* correspondentes nos dois *patches* comparados (BUADES; COLL; MOREL, 2005a).

Dada uma imagem de entrada  $u$ , o valor para um determinado *pixel* na posição  $\mathbf{x}$  da imagem é calculado pela Equação 1:

$$\text{NLM}(u(\mathbf{x})) = \frac{1}{C(\mathbf{x})} \sum_{\mathbf{y} \in \mathcal{N}(\mathbf{x})} \exp\left(-\frac{|u(\mathbf{x} + \cdot) - u(\mathbf{y} + \cdot)|^2}{h^2}\right) u(\mathbf{y}), \quad (1)$$

em que  $u(\mathbf{x})$  representa a intensidade do *pixels* na posição espacial  $\mathbf{x}$ ,  $\mathcal{N}(\mathbf{x})$  é a vizinhança ao redor de  $\mathbf{x}$ , e  $C(\mathbf{x})$  é um fator de normalização dado por:

$$C(\mathbf{x}) = \sum_{\mathbf{y} \in \mathcal{N}(\mathbf{x})} e^{-\frac{|u(\mathbf{x} + \cdot) - u(\mathbf{y} + \cdot)|^2}{h^2}}. \quad (2)$$

Nesta equação,  $|u(\mathbf{x} + \cdot) - u(\mathbf{y} + \cdot)|^2$  calcula o quadrado da diferença entre as vizinhanças de  $\mathbf{x}$  e  $\mathbf{y}$ , refletindo a dissimilaridade entre essas regiões. O parâmetro  $h$  controla o grau de suavização: um valor maior de  $h$  faz com que uma gama mais ampla de *pixels* da vizinhança contribua significativamente para o novo valor do *pixel* central, enquanto um valor menor de  $h$  restringe essa influência a *pixels* muito semelhantes, o que resulta em uma imagem mais suave, mas com possível perda de detalhes (BUADES; COLL; MOREL, 2005b; BUADES; COLL; MOREL, 2005a).

O termo  $u(\mathbf{y})$  representa a intensidade do *pixel* na posição  $\mathbf{y}$ , e cada um desses valores é multiplicado por um peso baseado na similaridade entre o *patch* centrado em  $\mathbf{x}$  e o *patch* centrado em  $\mathbf{y}$ . Esses pesos são determinados pela função exponencial  $e^{-\frac{|u(\mathbf{x} + \cdot) - u(\mathbf{y} + \cdot)|^2}{h^2}}$ , que garante que *patches* mais semelhantes tenham maior influência no cálculo da nova intensidade.

O fator de normalização  $C(\mathbf{x})$  garante que a soma dos pesos seja igual a 1, mantendo a consistência do valor de intensidade dos *pixels*. Como resultado, o valor de cada *pixel* é uma média ponderada das intensidades dos *pixels* da vizinhança, com maior peso dado aos *pixels* cujos *patches* são mais semelhantes ao *patch* de referência.

### 2.2.5 *Bias Field*

Outro artefato que pode estar presente nas imagens de RM é aquele causado pelo fenômeno *bias field*, que consiste em uma variação suave de baixa frequência dos níveis de intensidade da imagem, causando uma distorção gradual e espacialmente variável na intensidade dos *pixels* de tecidos homogêneos (GISPERT et al., 2004; LI; GORE; DAVATZIKOS, 2014).

Este artefato pode ser causado tanto pelo corpo do paciente quanto pela falta de homogeneidade das bobinas emisoras e receptoras de RF ou do campo  $B_0$ , entre outros fatores (GISPERT et al., 2004). O *bias field* pode interferir substancialmente na quantificação volumétrica dos tecidos cerebrais, causando erros em medições quantitativas e em classificações de imagens de RM (GISPERT et al., 2004; LI; GORE; DAVATZIKOS, 2014).

#### 2.2.5.1 Correção de *bias field*

Para corrigir a heterogeneidade de intensidade devido ao *bias field*, pode-se utilizar o método N4-ITK (TUSTISON et al., 2010). Esse método é uma variante do algoritmo *Nonparametric Nonuniform Intensity Normalization* (N3) (SLED; ZIJDENBOS; EVANS, 1998) e considera que o *bias field* é um sinal multiplicativo que suaviza o sinal originalmente emitido pelo tecido do paciente (SLED; ZIJDENBOS; EVANS, 1998; TUSTISON et al., 2010), conforme Equação 3. Nessa equação,  $v$  é o sinal medido na posição  $\mathbf{x}$ ,  $u$  é o sinal verdadeiro do tecido,  $f$  é o sinal do *bias field* e  $n$  é o sinal de ruído, ou seja,

$$v(\mathbf{x}) = u(\mathbf{x})f(\mathbf{x}) + n(\mathbf{x}). \quad (3)$$

O N4-ITK foi projetado para melhorar a eficácia e o desempenho do N3. Dentre as melhorias incorporadas, está a substituição da estratégia de suavização *B-spline* (utilizada para estimar o sinal de *bias field*) por uma alternativa mais rápida e robusta (TUSTISON et al., 2010). Para isso, o método incorpora um esquema hierárquico modificado, em que a correção do *bias field* é realizada em diferentes níveis de resolução. Em cada escala, o algoritmo realiza a correção do *bias field* de forma iterativa e incorpora informações de confiança dos dados em cada escala, a partir de máscara de confiança, refinando gradualmente a sua estimativa (TUSTISON et al., 2010). Além disso, o método também incorpora um esquema de otimização iterativa para melhorar o desempenho de convergência, demonstrando ser um método rápido, robusto e eficiente para correção do *bias field*.

## 2.3 Extração de atributos de textura

A análise de textura utiliza abordagens matemáticas para detectar variações espaciais nos níveis de intensidade dos *voxels* de uma imagem, permitindo capturar mudanças sutis

e complexas nas imagens (ZAIN; MAHMUD; GAN, 2023). A perda de corpos celulares neuronais, que ocorre durante a progressão da DA, pode gerar um aumento de água livre no espaço extracelular, prolongando o tempo de relaxação T1 e, com isso, diminuindo a intensidade do sinal na RM ponderada em T1 (LEE; LEE; KIM, 2020). Outros fatores que podem resultar em mudanças sutis na intensidade do sinal incluem desmielinização, depósitos de ferro e alteração do conteúdo de água modulado pela densidade sináptica alterada (LEE; LEE; KIM, 2020). Tais mudanças podem gerar padrões na imagem de RM quantificáveis por medidas texturais.

### 2.3.1 *Gray Level Co-occurrence Matrix (GLCM)*

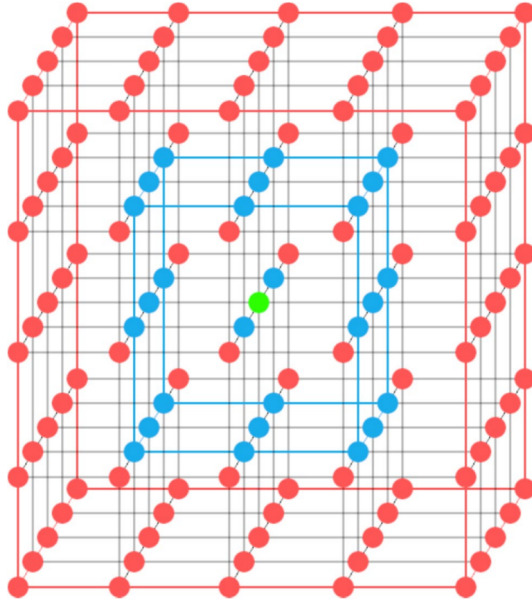
Diversos trabalhos identificaram que atributos de textura derivados da Matriz de Co-ocorrência de Níveis de Cinza (*Gray Level Co-occurrence Matrix*, GLCM) tem potencial discriminatório entre os grupos CN, CCL e DA (LEANDROU et al., 2020; SILVA et al., 2023; MALLIKA et al., 2023; SILVEIRA et al., 2024). Além disso, foi observado que mudanças nos depósitos da proteína tau podem refletir em alterações nos atributos de textura, destacando-se que esses atributos divergiram para cada região analisada (LEE; KIM; for the Alzheimer’s Disease Neuroimaging Initiative, 2021). Já Zhang et al. (2012) identificaram correlações entre atributos de textura derivados de GLCM com a pontuação obtida pela avaliação neuropsicológica Mini-Exame do Estado Mental (MEEM).

O GLCM é um método estatístico de segunda ordem que analisa a distribuição espacial de combinações de níveis de cinza em uma distância e orientações específicas (ZAIN; MAHMUD; GAN, 2023; SILVA et al., 2023). De forma geral, uma GLCM consiste em uma matriz simétrica quadrada que registra o número de vezes em que pares de *voxels* com determinados níveis de cinza ( $i, j$ ) aparecem juntos na imagem toda ou dentro de uma região de interesse (*Region of Interest*, ROI) (SILVEIRA et al., 2024).

No método isotrópico de GLCM 3D, a contagem considera todos os pares de *voxels* ( $i, j$ ) onde o *voxel*  $j$  está localizado em uma região cúbica isotrópica ao redor do *voxel*  $i$ , com uma distância de separação que pode variar de 1 a 5 *voxels* (SILVEIRA et al., 2024). Essa contagem é feita para 13 diferentes direções ao redor do *voxel* central, garantindo uma análise completa e tridimensional (ZAIN; MAHMUD; GAN, 2023). Nas bordas da ROI, os pares de *voxels* ( $i, j$ ) são contabilizados na matriz apenas se ambos os *voxels* estiverem inteiramente dentro da ROI (SILVEIRA et al., 2024). Uma representação esquemática da GLCM 3D com duas regiões cúbicas isotrópicas (representadas em azul e vermelho) pode ser vista na Figura 15, em que o ponto verde corresponde ao *voxel* de referência (*voxel*  $i$ ).

A Figura 16 mostra um exemplo de construção de uma matriz GLCM considerando uma imagem 2D de entrada, distância de 1 *pixel* e direção horizontal, considerando uma varredura da esquerda para direita (GLCM Oriental) e da direita para a esquerda (GLCM Ocidental) (FREIRE, 2019). As linhas e colunas laranjas representam os índices  $i$  e  $j$  da matriz, conforme explicado anteriormente.

Figura 15 – Representação esquemática da GLCM 3D.



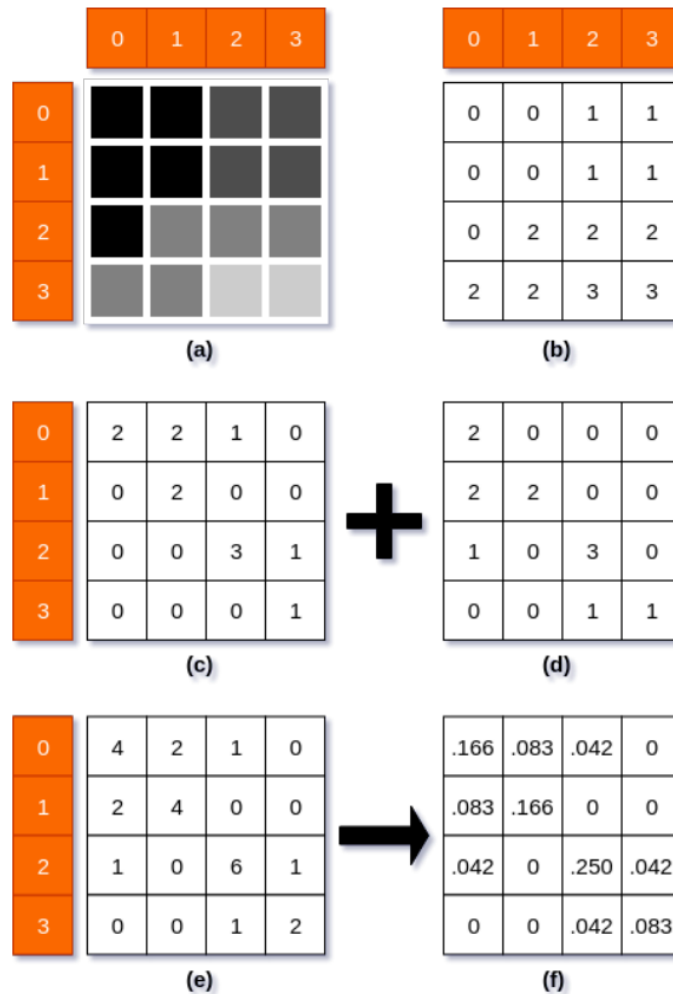
Fonte: Imagem extraída de Silveira et al. (2024).

Diante disso, as Figuras 16 (c) e (d) registram o número de vezes que um par de *pixels* (um com intensidade  $i$  seguido por outro com intensidade  $j$ ), são observadas a uma distância de 1 *pixel* entre os pares analisados da esquerda para direita (GLCM Oriental) e da direita para a esquerda (GLCM Ocidental), respectivamente. A Figura 16 (e) representa a relação espacial horizontal total, obtida a partir da soma das GLCM Oriental e Ocidental. A partir dessa matriz, normaliza-se os valores obtidos pela soma total dos elementos, obtendo-se assim a GLCM normalizada, mostrada na Figura 16 (f). A partir da GLCM normalizada, é possível obter vários atributos, como uniformidade, contraste, energia, correlação, variância, homogeneidade, entropia, média, variância, desvio padrão, entre outras (MALLIKA et al., 2023; SILVEIRA et al., 2024).

### 2.3.2 Redução de ruído para extração de atributos de textura

O trabalho de Zain, Mahmud e Gan (2023) identificou que o uso de métodos de redução de ruído antes da extração de atributos de textura pode afetar ligeiramente os valores desses atributos. Os autores compararam os atributos de contraste, correlação, energia e homogeneidade extraídos de GLCM a partir de imagens inteiras de pacientes CN e DA sem redução de ruído e com redução de ruído utilizando quatro métodos distintos (filtro de mediana, filtro de Wiener, filtro de equalização de histograma e filtro passa-baixa gaussiano). Apesar das imagens originais terem apresentados valores para os atributos de textura que permitiram distinguir melhor as classes CN e DA, o uso de filtros de redução de ruído é importante para as etapas de correção e segmentação, não testadas no trabalho citado.

Figura 16 – Exemplo da construção da matriz GLCM a partir de uma imagem 2D.



(a) Imagem de entrada, (b) Intensidades dos níveis de cinza da imagem de entrada, (c) GLCM Oriental, (d) GLCM Ocidental, (e) GLCM resultante, (f) GLCM normalizada. Fonte: Imagem extraída de Freire (2019).

Além disso, o trabalho não realizou experimentos com o filtro NLM (vide Subseção 2.2.4), não comparou métricas de regiões cerebrais específicas e também não verificou a distinção entre grupos de pacientes CCLp e CCLe. Essas lacunas destacam a necessidade de investigação adicional nesse campo, especialmente em relação à aplicação de diferentes métodos de redução de ruído e suas implicações para a classificação de imagens de RM.

### 2.3.3 Efeito do *scanner*

A intensidade reduzida do campo magnético em *scanners* de RM pode comprometer a detecção de alterações neurodegenerativas, dificultando, muitas vezes, a identificação de características relevantes para a diferenciação entre grupos clínicos distintos.

Isso foi evidenciado no estudo de Leandrou et al. (2020), no qual foram analisados os efeitos dos campos magnéticos de 1,5T e 3T utilizando imagens de RM ponderadas em T1 de 49 pacientes (14 CN, 14 CCL, 11 CCLp e 10 DA). Todos os pacientes foram

submetidos a exames de RM com ambos os campos magnéticos. Após a reamostragem das imagens para *voxels* isotrópicos de 1 mm<sup>3</sup> e a normalização da intensidade dos níveis de cinza pelo algoritmo N3, foram extraídos atributos de textura do hipocampo (baseados na técnica GLCM), espessura cortical e volumes dos hipocampos e das amígdalas.

Os atributos de textura incluíram o Momento Angular de Segunda Ordem (*Angular Second Moment*, ASM), contraste, correlação, variância, soma média, soma da variância e entropia, sendo as médias calculadas em quatro direções (0°, 45°, 90° e 135°), com uma distância unitária entre *pixels* adjacentes. Dessa forma, compararam-se as medidas de textura, volume e espessura entre as imagens adquiridas com 1,5T e 3T.

Nas imagens 1,5T, a entropia hipocampal mostrou diferença significativa apenas para CCL *vs* DA. Já os volumes dos hipocampos e das amígdalas mostraram diferenças significativas para CN *vs* DA. Em contraste, os atributos de textura do hipocampo extraídos de imagens 3T exibiram diferenças significativas em mais casos, como CN *vs* DA, CN *vs* CCL e CCL *vs* DA. Para a comparação CN *vs* DA, apenas a correlação não apresentou diferença significativa. Além disso, a espessura cortical foi significativamente diferente em CCL *vs* CCLp nas imagens 3T.

Ao comparar os dois campos magnéticos, foram encontradas diferenças estatisticamente significativas no ASM hipocampal, no volume da amígdala e na espessura cortical. Dentro dos grupos diagnósticos, diferenças significativas foram observadas apenas no grupo CN, especialmente para o ASM e no contraste dos hipocampos, além do volume das amígdalas. No caso dos atributos de textura, como correlação, variância, soma da variância e entropia, observaram-se diferenças significativas no grupo CN para esses quatro atributos. No grupo CCL, apenas a correlação apresentou diferença significativa, enquanto no grupo DA, diferenças significativas foram encontradas na variância e na soma das variâncias.

Esses achados podem ser explicados pelo fato de que imagens obtidas com *scanners* 3T possuem melhor SNR, facilitando a detecção de alterações neurodegenerativas. Além disso, os atributos de textura apresentaram mudanças estatisticamente mais significativas quando comparados aos volumes, o que sugere que alterações na textura hipocampal podem ocorrer antes de atrofia dos hipocampos e das amígdalas, corroborando os resultados de Lee, Lee e Kim (2020).

### 2.3.3.1 Correção do efeito do *scanner*

Os efeitos do *scanner* podem impactar significativamente a confiabilidade e a reprodutibilidade dos resultados da radiômica. Nesse contexto, Li et al. (2021) avaliaram como os métodos de pré-processamento, incluindo correção de *bias field* com N4 e reamostragem de imagens para 1mm isotrópico, e os métodos de harmonização, incluindo a normalização de intensidades aplicada às imagens de RM antes da extração dos atributos e o método ComBat aplicado diretamente nos atributos, influenciam na mitigação dos efeitos do *scan-*

ner e na reprodutibilidade das características radiômicas. Os métodos de normalização avaliados foram: Z-Score, *Fuzzy C-Means*, Modelo de Mistura de Gaussianas (*Gaussian Mixture Model*, GMM), *Kernel Density Estimate*, *WhiteStripe* e normalização de Nyúl. Em relação ao ComBat, foram analisadas três variantes aplicadas separadamente para cada classe de tecido (WM, GM, CC, etc.): ComBat padrão sem o uso de Bayesiano Empírico (BE), ComBat paramétrico usando BE paramétrico e ComBat não paramétrico usando BE não paramétrico, visando ajustes de dados mais robustos.

Para isso, foram selecionadas ROIs 2D de imagens de fantasmas homogêneos (9 ROIs para cada uma das 70 fatias escolhidas), de fantasmas heterogêneos (6 ROIs para cada uma das 30 fatias), de pacientes saudáveis (60 ROIs, sendo 20 de WM, 20 de GM e 20 do corpo caloso), e de pacientes com tumores cerebrais (6 ROIs das regiões saudáveis de WM), obtidas por *scanners* de 1,5T e 3T.

Após discretizar os níveis de cinza para 32 intervalos, foram extraídos 92 atributos radiômicos para cada ROI 2D, sendo 23 derivados da técnica GLCM, 16 da Matriz de Comprimento de Corrida de Níveis de Cinza (*Gray Level Run Length Matrix*, GLRLM), 16 da Matriz de Zona de Tamanho de Nível de Cinza (*Gray Level Size Zone Matrix*, GLSZM), 5 da Matriz de Diferença de Tonalidade de Níveis de Cinza Vizinhos (*Neighbouring Gray Tone Difference Matrix*, NGTDM), 14 de Matriz de Dependência de Níveis de Cinza (*Gray Level Dependence Matrix*, GLDM) e 18 de estatísticas de primeira ordem.

Os resultados mostraram que o uso do N4 não teve impacto significativo, indicando que as imagens analisadas não apresentavam um *bias field* elevado. A reamostragem preservou as características biológicas das imagens e, em alguns casos, reduziu oscilações nos histogramas, mas em outros, agravou-as devido ao ruído da interpolação. Além disso, a reamostragem preservou os valores médios de intensidade sem eliminar as diferenças significativas entre os *scanners* de 1,5T e 3T. Quanto aos métodos de normalização de intensidade, estes foram eficazes em remover as diferenças geradas por ajustes de *scanner*, como campo de visão (*field of view*, FOV) e dimensões da imagem, mas não eliminaram as diferenças entre os *scanners* de 1,5T e 3T.

A aplicação do método ComBat paramétrico mostrou-se eficaz para remover os efeitos do *scanner* e melhorar a reprodutibilidade dos atributos radiômicos, reduzindo as diferenças a zero ao comparar os *scanners* de 1,5T e 3T, independente do FOV, das dimensões das imagens e da aplicação prévia de N4 e/ou reamostragem. No entanto, as variâncias das características radiômicas permaneceram diferentes. Além disso, o ComBat paramétrico, quando combinado com qualquer método de normalização de intensidade, conseguiu remover os efeitos do *scanner* e preservar as informações biológicas. Os autores destacaram que a combinação de ComBat com as normalizações Z-Score e Nyúl foi particularmente eficaz, embora a relação relativa entre os atributos pudesse ser alterada pelo Z-Score.

Comparando as três variantes do método ComBat, o ComBat padrão demonstrou desempenho ligeiramente superior ao ComBat paramétrico e muito melhor do que o Com-

Bat não-paramétrico na ausência de pré-processamento e normalização de intensidade. Contudo, ao serem combinados com métodos de normalização de intensidade, todas as variantes do ComBat forneceram resultados satisfatórios, removendo os efeitos do *scanner* e preservando as informações biológicas. Assim, conclui-se que, embora os métodos de normalização de intensidade isoladamente não eliminem os efeitos do *scanner*, eles podem tornar os resultados de harmonização mais robustos, independentemente da variante do ComBat utilizada.

### 2.3.3.2 Método de harmonização ComBat

O ComBat foi originalmente desenvolvido para corrigir efeitos de lote em dados genéticos (JOHNSON; LI; RABINOVIC, 2007). Posteriormente, passou a ser aplicado também na análise de imagens de RM, com o objetivo de corrigir efeitos de *scanner*, permitindo que imagens adquiridas por diferentes *scanners* e/ou com configurações de aquisição distintas possam ser comparadas, mantendo-se as informações biológicas intactas.

O funcionamento básico do ComBat envolve a padronização dos dados, a estimação dos parâmetros de efeito de lote por meio de distribuições *a priori* empíricas paramétricas, e o ajuste desses dados para minimizar os efeitos de lote (JOHNSON; LI; RABINOVIC, 2007). Inicialmente, modela-se a intensidade da imagem ( $Y_{ijg}$ ) de acordo com a Equação 4 (JOHNSON; LI; RABINOVIC, 2007) a seguir:

$$Y_{ijg} = \alpha_g + X\beta_g + \gamma_{ig} + \delta_{ig}\varepsilon_{ijg}, \quad (4)$$

em que  $\alpha_g$  representa o efeito médio da intensidade do sinal associado à característica biológica  $g$ , comum a todas as imagens, independentemente do lote;  $X\beta_g$  é o efeito das covariáveis  $X$  (como idade, gênero, etc.) para a característica  $g$ , com  $\beta_g$  sendo o vetor de coeficientes dessas covariáveis;  $\gamma_{ig}$  é o efeito específico do lote  $i$  para a característica  $g$ ;  $\delta_{ig}$  é a variância associada ao lote  $i$  para a característica  $g$ , refletindo a variabilidade dos erros dentro do lote; e  $\varepsilon_{ijg}$  é o erro aleatório associado à imagem  $j$  no lote  $i$  para a característica  $g$ . Esse modelo, no entanto, não considera as variações entre as imagens dentro do mesmo lote, o que pode introduzir vieses nas estimativas dos efeitos de lote.

Para corrigir esses vieses, o ComBat ajusta os dados de cada lote ao estimar dois tipos de efeitos: efeitos de escala, que refletem a variabilidade da intensidade das imagens entre os lotes, e efeitos de localização, que capturam as diferenças nos níveis médios de intensidade entre os lotes. O método começa analisando imagens dentro de cada lote, utilizando a média e variância para padronizar os dados e remover as variações de escala e localização entre os lotes. A padronização é realizada pela Equação 5 (JOHNSON; LI; RABINOVIC, 2007) a seguir:

$$Z_{ijg} = \frac{Y_{ijg} - \hat{\alpha}_g - X\hat{\beta}_g}{\hat{\sigma}_g}, \quad (5)$$

sendo  $Z_{ijg}$  é o valor padronizado da característica  $g$  para a imagem  $j$  no lote  $i$ ,  $Y_{ijg}$  é o valor observado da característica  $g$ ,  $\hat{\alpha}_g$  é a estimativa do efeito médio da característica  $g$ , considerando a média de todas as imagens e todos os lotes,  $X\hat{\beta}_g$  representa o efeito das covariáveis  $X$  sobre a característica  $g$ , com  $\hat{\beta}_g$  representando os coeficientes estimados para essas covariáveis; e  $\hat{\sigma}_g$  é a estimativa da variância para a característica  $g$ , usada para padronizar os dados.

Com os dados padronizados, um modelo linear global é construído para integrar as variáveis biológicas e os efeitos ajustados de lote, levando em consideração as variações de escala e de localização, a partir dos dados de todos os lotes. Com base nesse modelo, estimam-se os efeitos aditivos (deslocamento da média) e multiplicativos (escala) para cada lote.

Na sequência, aplica-se a técnica de ajuste BE, que ajusta essas estimativas considerando a média e variância globais obtidas a partir do modelo linear global. Esse ajuste aproxima os efeitos de lotes individuais da média global estimada, levando em conta a variabilidade e o volume de dados de cada lote (JOHNSON; LI; RABINOVIC, 2007). As estimativas ajustadas dos efeitos de lote são então utilizadas para corrigir cada imagem, removendo as variações decorrentes do lote de origem. Esse processo resulta em dados mais consistentes e comparáveis entre os diferentes lotes.

## 2.4 Classificação supervisionada

O objetivo da identificação de quais pacientes CCL vão progredir ou não para DA pode ser vista como uma tarefa de classificação. Há muitas abordagens que podem ser usadas para esse fim, entretanto, para o contexto de captura de características para entender a progressão da DA é fundamental a escolha de um algoritmo de classificação que considere a relação entre as variáveis e não só as variáveis individualmente. Além disso, o modelo a ser utilizado deve ser capaz de extrair relações entre muitas variáveis diante de um conjunto de amostras pequeno.

Para isso, pode-se utilizar Florestas Aleatórias, algoritmo de aprendizado supervisionado construído pela combinação de algoritmos de árvores de decisão. Esse algoritmo é muito eficiente ao lidar com conjuntos de dados de alta dimensão e com presença de multicolinearidade, além de fornecerem estimativas para identificar quais atributos são mais relevantes na tarefa de classificação (FACELI et al., 2011; de ALVARENGA JÚNIOR, 2018; MARTELETTO, 2022).

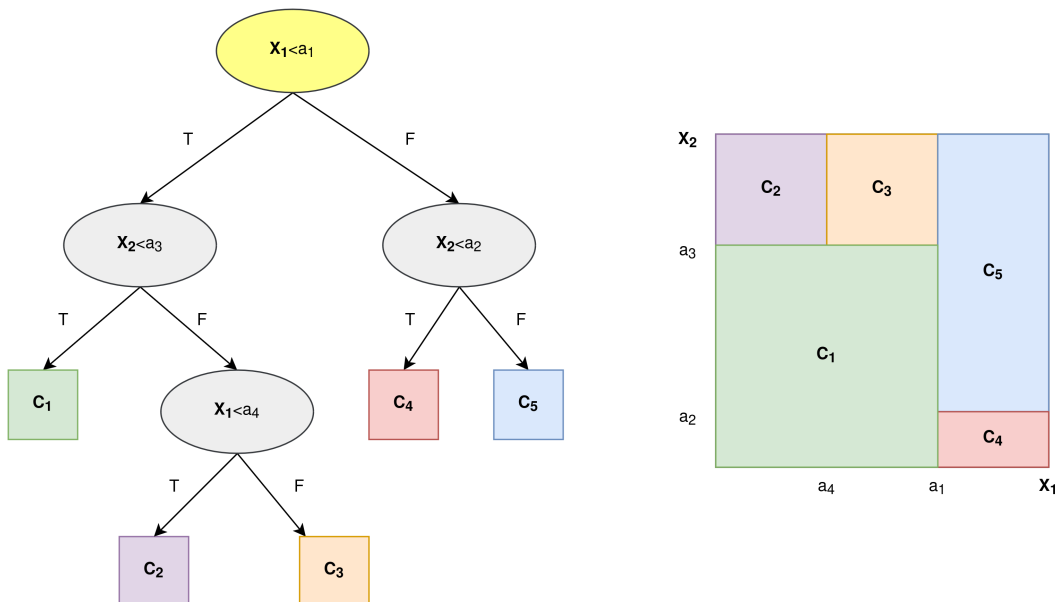
### 2.4.1 Árvore de Decisão

Uma árvore de decisão é um grafo acíclico direcionado, composto por uma raiz (nó sem nó-pai), nós internos (que possuem tanto nó-pai quanto nós-filhos) e nós-terminais

(ou folhas), que não possuem nós-filhos e atribuem a rotulação final à amostra de entrada (FACELI et al., 2011; de ALVARENGA JÚNIOR, 2018). Cada nó realiza um particionamento recursivo, baseado em testes condicionais aplicados aos atributos.

A Figura 17 apresenta uma árvore de decisão, juntamente com as regiões correspondentes a cada nó no espaço definido pelos atributos  $X_1$  e  $X_2$ . Nessa figura, a raiz está representada em amarelo, os nós internos em cinza e as folhas em formato quadrado. Cada ramo, representando o percurso entre a raiz e uma folha, contém uma cadeia de testes condicionais que envolvem operadores ( $=, \leq, \geq, \text{etc}$ ), verificando se o atributo atende a uma condição específica, atribuindo a classe final da amostra no final do ramo com base nas condições atendidas (FACELI et al., 2011; de ALVARENGA JÚNIOR, 2018).

Figura 17 – Árvore de decisão e as regiões de decisão no espaço de objetos.



Fonte: Imagem modificada de Faceli et al. (2011).

A escolha do atributo a ser utilizado em cada nó é baseada em medidas de impureza, como Entropia, Índice de Gini e a Razão de Ganho de Informação, selecionando o atributo que melhor discrimina as classes. Para isso, avaliam-se as probabilidades de que um exemplo pertença a cada um dos subconjuntos, calculando a proporção total de exemplos que é alocada a cada subconjunto (FACELI et al., 2011). Assim, se em um determinado nó todos os exemplos pertencerem à mesma classe, a impureza será nula. Por outro lado, se todas as classes estiverem presentes na mesma proporção naquele nó, a impureza será máxima (MARTELETTTO, 2022).

Em relação à Entropia, essa mede a aleatoriedade do atributo alvo, ou seja, quantifica o grau de incerteza em um nó em relação à classe à qual um objeto pertence, podendo ser calculada de acordo com a Equação 6:

$$H(p, q) = -p \log(p) - q \log(q), \quad (6)$$

onde  $p$  e  $q$  são as proporções de cada uma das duas classes em um nó. Assim, se todas as amostras no nó pertencerem a uma única classe ( $p = 0$  ou  $q = 0$ ), a entropia será zero. E quando as classes estão equilibradas ( $p = q$ ), a entropia será máxima, indicando que o nó tem pouca utilidade para o modelo (FACELI et al., 2011).

Entretanto, muitas vezes a entropia não apresenta valores extremos, tornando necessário calcular o Ganho de Informação para cada atributo. Para isso, seleciona-se o atributo que apresenta a maior diferença entre a entropia do nó-pai e a média ponderada das entropias dos nós descendentes (MARTELETTO, 2022). Isso requer, inicialmente, a medição da entropia esperada após a divisão das amostras com base em um atributo  $A$ , conforme a Equação 7:

$$E(A, p, q) = \sum_{i=1}^v \frac{p_i + q_i}{p + q} H(p_i, q_i), \quad (7)$$

em que  $p_i$  e  $q_i$  são as proporções das duas classes no  $i$ -ésimo subconjunto após a divisão (FACELI et al., 2011). O valor resultante reflete a impureza média dos subconjuntos resultantes da divisão, servindo como uma medida da incerteza desses subconjuntos.

Por fim, para avaliar a quantidade de informação adquirida ao dividir as amostras de acordo com um teste condicional para o atributo  $A$ , mede-se o Ganho de Informação de acordo com a Equação 8:

$$IG(A, p, q) = H(p, q) - E(A, p, q). \quad (8)$$

Um valor alto de  $IG$  indica que a divisão com o atributo  $A$  resulta em subconjuntos mais homogêneos (ou seja, com menor impureza), o que é desejável ao construir uma árvore de decisão. Se o  $IG$  for zero, isso indica que a divisão não trouxe informação adicional, não tendo utilidade ao modelo.

Já o Índice de Gini mede a variância total entre as classes em um nó específico, sendo uma medida da probabilidade de erro de classificação (de ALVARENGA JÚNIOR, 2018), conforme Equação 9:

$$i(t) = 1 - \sum_i p_i^2, \quad (9)$$

em que  $p_i$  é a proporção de cada classe  $i$  no nó avaliado. O valor resultante varia de 0 a 1, onde 0 indica pureza perfeita (apenas uma classe) e 1 indica a máxima impureza (classes uniformemente distribuídas).

Também pode-se utilizar a Entropia Cruzada, que mede a diferença entre duas distribuições de probabilidade, conforme a Equação 10:

$$H_{cruzada}(p, q) = -\sum_i p_i \log(q_i). \quad (10)$$

Essa medida compara uma distribuição verdadeira  $p$  com uma distribuição predita  $q$  para cada classe  $i$  e é utilizada para verificar quão bem as previsões de um modelo correspondem aos rótulos verdadeiros.

Além dessas medidas, também é fundamental controlar a extensão das árvores, uma vez que árvores com muitas divisões tem maior custo computacional e podem levar ao sobreajuste (de ALVARENGA JÚNIOR, 2018; MARTELETTO, 2022). Para isso, ajusta-se parâmetros que regulam o número máximo de atributos a serem considerados para formar os nós e a profundidade máxima da árvore (MARTELETTO, 2022). Além disso, com a árvore já construída, pode-se realizar o processo de poda, que consiste em eliminar nós ou ramos que são pouco significativos.

Embora o modelo de árvore de decisão seja capaz de capturar relações complexas em um problema, ele geralmente não é competitivo em comparação com as melhores técnicas de aprendizado supervisionado, devido à sua alta variância (de ALVARENGA JÚNIOR, 2018). No entanto, ao utilizar múltiplas árvores de decisão em conjunto, como no caso das Florestas Aleatórias, esse problema de variância é mitigado.

## 2.4.2 Florestas Aleatórias

As Florestas Aleatórias são modelos do tipo *ensemble* que combinam as respostas individuais de múltiplas árvores de decisão, selecionando o resultado mais votado como a previsão final do modelo (MARTELETTO, 2022; de ALVARENGA JÚNIOR, 2018). Para alcançar isso, diferentes subconjuntos de preditores são usados na construção de cada árvore, sem a necessidade de que cada preditor seja relevante por si só. Isso resulta em árvores menos ajustadas individualmente, mas que, quando combinadas, proporcionam um desempenho mais robusto (FACELI et al., 2011).

Ao considerar combinações variadas de preditores em cada árvore, esses modelos conseguem explorar um espaço maior de atributos e suas interações (FACELI et al., 2011). Isso significa que elas têm uma chance maior de identificar padrões complexos e sutis nos dados que poderiam passar despercebidos em abordagens mais tradicionais. O efeito da "votação" entre as árvores reduz a variância e aumenta a robustez do modelo, mitigando o impacto de *outliers* ou de padrões não representativos nos dados de treinamento (de ALVARENGA JÚNIOR, 2018).

A construção das Florestas Aleatórias envolve a realização de uma amostragem aleatória com reposição (também chamada de *Bootstrap*), a partir do conjunto de dados de treino, para formar novos conjuntos de instâncias (de ALVARENGA JÚNIOR, 2018). Cada um desses subconjuntos é utilizado para desenvolver uma árvore de decisão, construída de forma independente e com profundidade variável, sem poda (MARTELETTO, 2022). Para formar cada nó, um subconjunto aleatório com  $m$  atributos é selecionado, e o atributo que fornecer a melhor divisão é escolhido (MARTELETTO, 2022).

As árvores independentes são então combinadas, caracterizando o modelo como um método de *Bagging* (também chamada de *Bootstrap Aggregating*) (de ALVARENGA JÚNIOR, 2018; MARTELETTO, 2022). Cada árvore contribui igualmente para a previsão final, e o resultado é determinado pela maioria das votações entre as árvores. Com isso,

há redução da variância e o modelo tende a ter um desempenho superior em relação às árvores de decisão individuais (de ALVARENGA JÚNIOR, 2018).

Diante disso, a saída do modelo pode ser descrita conforma a Equação 11:

$$\hat{f}_{fa}^B(x_i) = \frac{1}{B} \sum_{b=1}^B T(x_i, \Theta_b), \quad (11)$$

em que,  $B$  é o total de árvores,  $T()$  é a resposta de uma árvore  $b$ , com parâmetros  $\Theta_b$ , para um vetor de entrada  $x_i$  (de ALVARENGA JÚNIOR, 2018).

Para a construção de uma Floresta Aleatória, é fundamental ajustar o hiperparâmetro  $B$ , controlando o número de árvores de decisão presentes no modelo. Um número elevado de árvores pode levar ao maior custo computacional sem necessariamente apresentar uma melhora no desempenho (de ALVARENGA JÚNIOR, 2018).

Além disso, deve-se controlar o número de atributos a serem considerados para o particionamento em cada nó (parâmetro  $m$ ), pois isso ajusta o grau a correlação entre as árvores da Floresta Aleatória, proporcionando uma maior capacidade de generalização ao modelo. Na presença de um número elevado de atributos no conjunto de dados e com a maioria sendo irrelevante para a predição, um pequeno valor de  $m$  pode levar a um desempenho ruim do modelo (de ALVARENGA JÚNIOR, 2018). No contexto de classificação, o valor de  $\sqrt{p}$  é comumente usado para esse hiperparâmetro, com  $p$  sendo o número total de atributos no conjunto de dados (de ALVARENGA JÚNIOR, 2018).

Na Tabela 5, são sumarizados os hiperparâmetros e parâmetros importantes para o ajuste de modelos de Floresta Aleatória.

Tabela 5 – Hiperparâmetros e parâmetros de modelos de Floresta Aleatória.

Parâmetros	Hiperparâmetros
- Atributo utilizado em cada nó (de cada árvore)	- Número de árvores ( $B$ )
- Valor de corte em cada nó (de cada árvore)	- Profundidade das árvores
	- Quantidade de atributos no particionamento ( $m$ )
	- Critério de pureza

Fonte: Tabela modificada de Alvarenga Júnior (2018).

### 2.4.2.1 Importância dos atributos

As Florestas Aleatórias possuem como vantagem a fácil interpretabilidade sobre quais atributos são mais relevantes para a tarefa de predição. Para isso, primeiramente, calcula-se a importância dos nós individuais por árvore, conforme mostrado na Equação 12:

$$ni_j = w_j C_j - w_{esquerda(j)} C_{esquerda(j)} - w_{direita(j)} C_{direita(j)}, \quad (12)$$

onde,  $ni_j$  representa a importância do nó  $j$ , com  $w_j$  sendo o peso das amostras que atingem o nó  $j$ ,  $C_j$  o valor de impureza do nó  $j$ ,  $w_{esquerda(j)}$  e  $C_{esquerda(j)}$  sendo o peso e o valor

de impureza do nó-filho à esquerda de  $j$ , e  $w_{direita(j)}$  e  $C_{direita(j)}$  sendo o peso e o valor de impureza do nó-filho à direita de  $j$  (VELAZQUEZ; LEE, 2021). Com isso, obtém-se a importância de um nó em termos de como ele melhora a pureza das amostras. A diferença entre a impureza do nó-pai e a soma das impurezas dos nós-filhos fornece uma medida de quanto o nó-pai contribui para a separação das classes (ou redução do erro).

Uma vez que a importância de cada nó está determinada, a importância da variável  $i$  por árvore ( $fi_i$ ) pode ser calculada, dividindo-se a soma das importâncias dos nós que usam essa variável por todas as importâncias dos nós da árvore, fornecendo uma medida relativa da contribuição da variável para a árvore em relação à contribuição total, conforme a Equação 13:

$$fi_i = \frac{\sum_{j \in \{\text{nós} | \text{nó } j \text{ realiza uma divisão com a característica } i\}} ni_j}{\sum_{k \in \text{todos os nós}} ni_k}. \quad (13)$$

Em seguida, a importância da variável  $i$  deve ser normalizada em relação à soma das importâncias de todas as variáveis, conforme Equação 14:

$$normfi_i = \frac{fi_i}{\sum_{j \in \text{todos os atributos}} fi_j}. \quad (14)$$

Isso resulta em uma escala de 0 a 1, fazendo com que o valor normalizado ( $normfi_i$ ) seja interpretado mais facilmente, permitindo comparações diretas entre variáveis (VELAZQUEZ; LEE, 2021).

Por fim, calcula-se a importância média da variável  $i$  em toda a Floresta Aleatória, somando as importâncias normalizadas em todas as árvores e dividindo pelo número total de árvores, conforme Equação 15:

$$RFfi_i = \frac{\sum_{j \in \text{todas as árvores}} normfi_{ij}}{T}. \quad (15)$$

Com isso, é possível obter uma medida que reflete a contribuição da variável para a predição em todo o conjunto de árvores do modelo, permitindo identificar quais atributos estão envolvidos na progressão da DA.

## 2.5 Considerações finais

Este capítulo apresentou informações complementares essenciais para o entendimento deste estudo, abordando as causas, o diagnóstico e o tratamento da DA, além de explorar o processo de formação e aquisição de imagens de RM e os desafios associados ao processamento dessas imagens para a extração de atributos radiômicos. No próximo capítulo, serão discutidos estudos já publicados que representam o estado da arte relacionado a este trabalho.



---

## Capítulo 3

# Revisão Bibliográfica

---

*Este capítulo apresenta uma revisão das diversas técnicas propostas na literatura para a extração de características de imagens de RM baseada em engenharia de atributos visando a distinção entre CCLp e CCLe para auxiliar no diagnóstico da doença de Alzheimer.*

### 3.1 Trabalhos correlatos com engenharia de atributos

No estudo de Shu et al. (2021), um modelo radiômico integrado foi desenvolvido e validado utilizando imagens de RM de todo o cérebro para prever a conversão de CCL para DA. A partir do banco de dados da ADNI, foram analisados 357 pacientes com CCL, dos quais 154 progrediram para DA em até 48 meses de acompanhamento. As imagens de RM ponderadas em T1 foram segmentadas automaticamente em GM, WM e CSF, e foram extraídas 378 características radiômicas de cada tecido, resultando em 1134 características por imagem.

A partir do grupo de treinamento, uma assinatura radiômica foi construída utilizando regressão logística, após a redução de dimensionalidade com os algoritmos *minimum Redundancy Maximum Relevance* (mRMR) e *Least Absolute Shrinkage and Selection Operator* (LASSO). Para extrair um conjunto diversificado de características radiômicas, utilizou-se informações do histograma, características de Haralick (baseadas em GLCM), GLRLM, que quantifica a distribuição dos comprimentos de regiões homogêneas na imagem, e GLSZM, que avalia as propriedades de tamanho das regiões de diferentes níveis de cinza na imagem. Essas características foram avaliadas com base em sua robustez, determinada pelo teste de correlação de Spearman, sendo consideradas robustas aquelas com coeficiente de correlação superior a 0,8.

A análise de regressão logística *stepwise* revelou que a variante APOE4 (uma variante do gene APOE), as escalas neuropsicológicas e a assinatura radiômica foram preditores independentes da progressão de CCL para DA. Para construir o modelo integrado, foram utilizados métodos como seleção retroativa e regressão múltipla, com base em critérios como o Critério de Informação de Akaike (*Akaike's Information Criterion*, AIC) para a escolha de preditores e o Fator de Inflação da Variância para mitigar multicolinearidade. Algoritmos de aprendizado de máquina, como Máquinas de Vetores de Suporte (*Support Vector Machines*, SVM), foram selecionados pela sua robustez em reconhecer padrões complexos de neurodegeneração.

A área sob a curva *Receiver Operating Characteristic* (ROC), conhecida como *Area Under the Curve* (AUC) para o modelo no conjunto de teste foi de 0,807, com especificidade de 73,8% e sensibilidade de 74,5%. Além disso, o modelo mostrou-se mais eficaz na previsão da progressão de CCL para DA em até 12 meses, com AUC de 0,814, sensibilidade de 72,6% e especificidade de 79,8%.

Usando apenas a assinatura radiômica para a predição, o modelo obteve uma acurácia de 69,2% no conjunto de teste, com especificidade de 72,1% e sensibilidade de 65,9%. Os autores concluem que o modelo integrado, baseado em radiômica de todo o cérebro, apresenta desempenho superior em relação aos preditores independentes, identificando e prevendo com precisão os pacientes com CCL em alto risco de progressão para DA, destacando a importância dos biomarcadores radiômicos na fase inicial da doença.

Já no estudo de Calandrelli et al. (2022), foi avaliada a precisão das análises quantitativa e semi-quantitativa na detecção de padrões de atrofia nas regiões do lobo temporal médio (LTM), temporal anterior (TA) e fronto insular (FI) para diferenciar entre CCL amnésico estável (aCCLe) e progressivo (aCCLp) em um acompanhamento de 24 meses. Foram utilizadas imagens de RM da linha de base de 133 pacientes encaminhados para a *Memory Clinic of the Policlinico Universitario "A. Gemelli"* (MC-PUAG) em Roma, dos quais 54 eram aCCLe e 79 aCCLp.

Na análise semi-quantitativa, neurorradiologistas pontuaram as regiões em cada imagem com base em escalas visuais que refletem o grau de atrofia. Para a análise quantitativa, foi utilizado o *Freesurfer* para extrair volumes subcortical dos hipocampos e obter volumes e espessuras corticais das regiões, utilizando parcelamentos corticais conforme o atlas de Desikan. As pontuações visuais foram então comparadas à espessura cortical média e aos volumes corticais normalizados das estruturas da mesma região, selecionando apenas as diferenças estatisticamente significativas entre os subgrupos para a classificação.

Ao comparar os subgrupos de aCCL, foram detectadas diferenças significativas nas pontuações das escalas de atrofia do LTM (*medial temporal lobe atrophy*, MTA), TA e FI, na espessura cortical dessas regiões, e nos volumes de matéria cinzenta do LTM e TA, com escores mais altos e menores medidas de espessura/volume em aCCLp, do que em aCCLe. Conforme a Tabela 6, as pontuações visuais para MTA, TA e FI demonstraram elevado

valor diagnóstico, especialmente a pontuação de TA, que diferenciou aCCLe de aCCLp com a melhor combinação de sensibilidade (80%) e especificidade (99%). Em contraste, as medidas quantitativas mostraram precisão diagnóstica baixa para espessura cortical e moderada para volumes.

A análise multivariada para pontuações visuais combinadas mostrou uma melhoria significativa na diferenciação entre aCCLe e aCCLp (AUC = 0,85; sensibilidade = 95%). Além disso, a combinação de medidas de espessura cortical e volume do hipocampo também apresentou bom desempenho (AUC = 0,77; sensibilidade = 76%). Em contraste, ao considerar apenas medidas combinadas de espessura, o valor diagnóstico foi semelhante ao de uma única região (AUC = 0,62; sensibilidade = 57%), e as medidas combinadas de volume mostraram valor diagnóstico similar (AUC = 0,74), mas com uma sensibilidade inferior (61%).

Tabela 6 – Principais resultados do trabalho de Calandrelli et al. (2022).

Variáveis	AUC	Sens. (%)	Espec. (%)
Pontuações visuais de MTA	0,83	56,0	83,0
Pontuações visuais de TA	0,80	80,0	<b>99,0</b>
Pontuações visuais de FI	0,80	57,0	85,0
Combinação das pontuações visuais de MTA, TA e FI	<b>0,85</b>	<b>95,0</b>	64,0
Média da espessura cortical de LTM	0,59	60,0	53,0
Média da espessura cortical de TA	0,61	61,0	59,0
Média da espessura cortical de FI	0,58	62,5	46,0
Combinação das médias das espessuras corticais de LTM, TA e FI	0,62	57,0	75,0
Volumes normalizados de LTM	0,73	71,0	65,0
Volumes normalizados de TA	0,69	66,0	60,0
Combinação dos volumes normalizados de LTM e TA	0,74	61,0	77,0
Combinação das médias das espessura corticais de LTM, TA e FI com volume hipocampal	<b>0,77</b>	<b>76,0</b>	67,0

Fonte: Informações retiradas de Calandrelli et al. (2022).

No trabalho de Xie et al. (2023), foram analisadas medidas estruturais, biomarcadores plasmáticos e suas combinações para identificar quais poderiam prever a atrofia estrutural longitudinal e o declínio cognitivo relacionados à DA. Para isso, foram utilizadas informações de 245 CN e 361 CCL da base ADNI (ADNI-GO e ADNI-2), incluindo biomarcadores plasmáticos como proteína Tau fosforilada na posição 181 (p-Tau<sub>181</sub>) e neurofilamento leve (NFL), além de imagens de RM estrutural ponderadas em T1.

As medidas estruturais extraídas incluíram o volume do hipocampo anterior e posterior, a espessura do Córtex Entorrinal (CER) (sub-região cortical do LTM), as espessuras das Áreas de Brodmann (AB) 35 e 36 (AB35 e AB36), o volume do Córtex Parahipocampal (*Parahippocampal Cortex*, PHC) e o volume intracraniano (*Intracranial Volume*,

ICV). Os grupos CN e CCL foram subdivididos com base na presença de biomarcadores de  $\beta$ -amiloide (BetaA) (presente: BetaA+; ausente: BetaA-).

Utilizando dados longitudinais, como a mudança do volume da AB35 (uma das melhores medidas para discriminar pacientes nas fases iniciais da DA dos controles BetaA- (XIE et al., 2020)), a mudança na ADNI-MEM (medida derivada de um subconjunto de testes de memória disponibilizada pela ADNI), e a pontuação do *Clinical Dementia Rating - Sum of Boxes* (CDR-SOB) (teste que avalia diferentes domínios cognitivos), foram desenvolvidos modelos para prever essas alterações ao longo do tempo. Posteriormente, para identificar quais biomarcadores seriam preditivos para cada medida longitudinal, aplicou-se o método *stepwise* em modelos lineares mistos, utilizando critérios como o AIC. Os modelos base incluíam apenas variáveis demográficas (como idade, sexo e educação), ICV e status do alelo APOE4. Medidas estruturais de imagens de RM e biomarcadores plasmáticos foram adicionados iterativamente, com cada medida sendo mantida no modelo apenas se melhorasse significativamente o AIC, até que nenhuma melhoria adicional fosse obtida.

No total, 15 modelos finais foram gerados, sendo seis para a mudança de volume da AB35 (um para cada grupo: CN, CN BetaA-, CN BetaA+, CCL, CCL BetaA- e CCL BetaA+), seis para a mudança do ADNI-MEM e três para a mudança do CDR-SOB (um para cada grupo: CCL, CCL BetaA- e CCL BetaA+). Os biomarcadores de plasma e de RM estrutural da linha de base foram ambos incluídos na maioria dos modelos finais (9 dos 15 modelos), apoiando a hipótese de que esses biomarcadores fornecem informações complementares na previsão da progressão da DA.

Adicionalmente, foi realizada uma análise de regressão logística para avaliar a capacidade dos modelos em discriminar entre dois grupos: progressores rápidos (33% com maior progressão longitudinal) e lentos (33% com menor progressão). Os resultados revelaram que a combinação dos biomarcadores plasmáticos e RM alcançou AUCs entre 0,89 e 0,93 na discriminação desses dois grupos no conjunto CCL, enquanto no subgrupo CCL BetaA+, os AUCs variaram entre 0,85 e 0,88. Esses resultados proporcionam um critério confiável para diferenciar entre indivíduos com diferentes taxas de progressão da doença. Além disso, os biomarcadores combinados consistentemente superaram os biomarcadores individuais na maioria das análises, destacando a complementaridade entre os dois tipos de biomarcadores.

Como apresentado na Tabela 7, o volume do hipocampo posterior foi consistentemente incluído em todos os modelos relacionados ao grupo CCL BetaA+, enquanto que, nos modelos do grupo CCL BetaA-, a espessura do AB35 se destacou como o biomarcador mais preditivo, sendo incluída em dois modelos, seguida pela espessura do CER, incluída em um modelo. A variação na inclusão de diferentes combinações de biomarcadores entre os grupos CCL BetaA- e CCL BetaA+ sugere possíveis diferenças na patologia subjacente e na progressão da doença. Considerando que indivíduos com CCL BetaA+ apresentam

maior risco de progressão para DA em comparação com aqueles com CCL BetaA–, essa capacidade preditiva é importante para entender como esses biomarcadores podem ser úteis na prática clínica para identificar pacientes em risco de progressão da doença.

Tabela 7 – Principais resultados do trabalho de Xie et al. (2023) para os grupos CCL BetaA– e CCL BetaA+.

Variáveis dependentes	Grupo	Medidas incluídas no modelo	AUC
Mudança no volume de AB35	CCL BetaA–	Espessura da AB35	<b>0,85</b>
	CCL BetaA+	Volume do hipocampo posterior; p-Tau <sub>181</sub>	0,85
Mudança no ADNI-MEM	CCL BetaA–	Espessura da AB35; NfL	0,80
	CCL BetaA+	Volume do hipocampo posterior; Espessura da AB36; p-Tau <sub>181</sub>	<b>0,88</b>
Mudança no CDR-SOB	CCL BetaA–	Espessura do CER; NfL	0,78
	CCL BetaA+	Volume do hipocampo posterior; Espessura da AB35; p-Tau <sub>181</sub>	0,86

Fonte: Informações retiradas de Xie et al. (2023).

Embora as medidas de RM estrutural tenham sido incluídas na maioria dos modelos, houve uma variação significativa nos biomarcadores mais preditivos: a espessura do AB35 em 8 modelos, o volume do hipocampo posterior em 6 modelos, o volume do hipocampo anterior em 2 modelos, a espessura do CER em 1 modelo e a espessura do AB36 em 1 modelo. Essa diversidade nos biomarcadores mais preditivos torna desafiador identificar quais regiões específicas do LTM são mais cruciais para a progressão da doença. Portanto, é fundamental considerar todas as informações disponíveis sobre espessura, volume e outras características das diferentes regiões dentro do LTM para obter medições mais precisas e confiáveis das mudanças associadas à progressão da doença.

No estudo de Xiao et al. (2021), foi empregado um modelo de regressão logística esparsa com *elastic net* generalizada, que combina a regularização  $L_p$  e a regularização  $L_2$ , para o diagnóstico precoce da DA. Essas técnicas de regularização possibilitam gerar soluções esparsas, o que significa que apenas os atributos mais relevantes são mantidos no modelo, tornando-o mais simples. Além disso, essas técnicas garantem que as regiões cerebrais correlacionadas sejam incluídas simultaneamente na análise, melhorando a precisão do diagnóstico.

Os atributos do modelo foram compostos pelos volumes de GM de 90 ROIs, obtidos após o registro das imagens de cada paciente com o atlas anatômico *Automated Anatomical*

*Labeling.* As ROIs mais relevantes foram selecionadas usando o método de regularização *elastic net* generalizada. Este processo utilizou imagens de RM estrutural da base de dados ADNI, que incluiu 45 pacientes CCLe e 51 CCLp.

Para avaliar o desempenho de classificação entre CCLp e CCLe, aplicou-se um processo de validação cruzada aninhada em dois níveis. No nível externo, o conjunto de dados foi dividido em 10 partes, reservando uma para teste em cada iteração. No nível interno, cada parte do conjunto de treinamento foi submetida a uma validação cruzada de 10 partes para ajustar e selecionar os melhores parâmetros  $\lambda_1$  e  $\lambda_2$ . O melhor resultado foi obtido com o método de regularização  $L_p + L_2(p = 0.8)$ , com o qual alcançou-se uma acurácia de 75,87%, sensibilidade de 73,71% e especificidade de 77,51%. Além disso, as 10 principais regiões cerebrais discriminativas identificadas pelo método proposto incluem o hipocampo esquerdo, amígdala esquerda, giro temporal médio esquerdo, *precuneus*, giro parahipocampal direito, giro occipital superior esquerdo, giro pré-central direito, insula esquerda, angular esquerda e cúneo esquerdo, as quais podem ser biomarcadores potenciais de imagem de RM para diagnóstico da DA.

Já Minhas et al. (2021) desenvolveram um método de predição de diagnósticos empregando a informação da série temporal para a classificação de CCLp e CCLe usando um classificador SVM, selecionado para a classificação binária devido à disponibilidade limitada de dados e à robustez deste classificador. Para treinar e validar o modelo, foram usadas 8 medidas neuropsicológicas (MN) e 249 medidas derivadas de imagens de RM.

As MNs utilizadas foram: ADAS 13, Teste Auditivo Verbal de Rey (TAVR), Teste do Desenho do Relógio, Teste de Cópia do Relógio, Pontuação Total de Recordação Imediata, MEEM, Teste de Trilhas A e de Trilhas B. Já os atributos das imagens de RM consistiram em volumes de regiões cerebrais obtidos após o parcelamento cortical e o parcelamento da WM, a área de superfície das regiões cerebrais e a espessura cortical de regiões cerebrais fornecidas pela ADNI.

Para avaliar a significância dos biomarcadores, foi realizado teste t de *Student* com amostras das leituras de biomarcadores na linha de base. Em seguida, os biomarcadores foram classificados de acordo com seu impacto na discriminação entre CCLp e CCLe. Por fim, para selecionar os atributos relevantes, utilizou-se um modelo preditivo para avaliar diferentes combinações de biomarcadores distintos, selecionando aqueles que melhoraram o desempenho do modelo.

A partir dos dados de treinamento, estimou-se a taxa de mudança de um preditor ao longo do tempo e sua média na classe. Em seguida, determinou-se os pesos dos deslocamentos do gradiente usando um modelo de regressão linear, definindo quanto cada deslocamento do gradiente de diferentes preditores contribui para prever o próximo valor de um preditor específico. Com isso, calculou-se a média dos pesos utilizando a validação cruzada e desenvolveu-se um método para prever as próximas leituras dos preditores com base em duas medidas consecutivas anteriores dos preditores. Após isso, utilizou-se a

informação da série temporal (sequência de leituras dos preditores ao longo do tempo) para criar um modelo de classificação dos pacientes em CCLp e CCLe.

Para treinar e validar o modelo desenvolvido, foram utilizadas informações de pacientes CCL disponíveis na base de dados ADNI-1 que possuíam, no mínimo, três leituras consecutivas de acompanhamento. Pacientes que converteram para DA antes do 36<sup>o</sup> mês foram classificados como CCLp. Na tarefa de predição com 6 meses de antecedência, foram considerados indivíduos com três leituras consecutivas de acompanhamento semestral, totalizando 49 CCLp e 70 CCLe. Para a previsão com 1 ano de antecedência, foram incluídos indivíduos com três leituras anuais consecutivas, resultando em 35 CCLp e 50 CCLe.

Das 257 preditores analisados, 98 deles demonstraram ter uma influência relevante na capacidade do modelo de prever com precisão as leituras futuras. Os atributos mais significativos para a predição da conversão com antecedência de 6 meses foram TAVR, ADAS, volume da amígdalas (esquerda e direita) e volume do entorrinal direito. Já para a predição com 1 ano de antecedência, os principais atributos foram TAVR, ADAS, espessura cortical do istmo do giro do cíngulo (esquerdo), espessura cortical média do fusiforme (esquerdo) e espessura cortical média do istmo do giro do cíngulo (direito). Observou-se que os volumes cerebrais têm maior relevância em previsões de curto prazo, enquanto a espessura cortical média é mais relevante para previsões de 1 ano.

Além disso, ao comparar a predição de leituras futuras em um ponto no tempo utilizando medidas de NM e RM separadamente e em combinação, observou-se que o erro de predição foi maior no modelo que utilizou apenas medidas de NM e menor no modelo que utilizou somente medidas de RM. Isso reforça a imprevisibilidade do comportamento humano e as dificuldades de modelá-lo, especialmente em períodos curtos de acompanhamento (6 meses). Também destaca-se a previsibilidade linear e a irreversibilidade da atrofia cerebral. A previsão de marcadores de RM com 6 meses de antecedência mostrou-se mais precisa do que com 1 ano de antecedência, uma vez que a atrofia cerebral é um processo relativamente lento, com mudanças menores em períodos de acompanhamento mais curtos.

Para predição dentro de 6 meses, o modelo que utilizou tanto medidas de NM quanto de RM (19 atributos) apresentou maior AUC (91,2%), seguido pelo modelo que utilizou apenas medidas de RM (13 atributos) (AUC = 88%) e pelo modelo que utilizou apenas NM (2 atributos) (AUC = 87,1%). Para a predição em um intervalo de 1 ano, o modelo que combinou medidas de NM e RM (9 atributos) também obteve a maior AUC, com 95,7%, seguido pelo modelo contendo apenas NM (4 atributos) (AUC = 91,5%) e pelo modelo que utilizou apenas medidas de RM (13 atributos) (AUC = 90,5%). Os resultados de acurácia, sensibilidade e especificidade de cada modelo desenvolvido são apresentados na Tabela 8. Nela, é possível observar que o modelo que utilizou medidas de NM e RM forneceu maior sensibilidade e especificidade em comparação com os modelos unimodais

para o tempo de predição de 6 meses, enquanto o modelo que utilizou somente medidas NM para tempo de predição de 1 ano apresentou melhor desempenho. Isso sugere que tratar cada tipo de atributo de forma independente, em vez de combiná-los em um único espaço, pode proporcionar uma análise mais detalhada e precisa, resultando em um desempenho superior na identificação de pacientes CCLp.

Tabela 8 – Principais resultados do trabalho de Minhas et al. (2021) para distinguir CCLp e CCLe.

Tempo de predição	Tipo de medidas utilizadas	Acurácia (%)	Sens.(%)	Espec.(%)
6 meses	NM	77	73,5	80
	RM	79,8	71,8	71,8
	NM + RM	<b>84,9</b>	<b>81,7</b>	<b>87,1</b>
1 ano	NM	<b>84,7</b>	82,9	86
	RM	83,5	74,28	<b>90</b>
	NM + RM	81	<b>85,7</b>	70

Fonte: Informações retiradas de Minhas et al. (2021).

No trabalho de Syaifullah et al. (2021), foi desenvolvido um *software* chamado *Brain Anatomical Analysis Using Diffeomorphic Deformation* para análise e diagnóstico da DA e previsão da progressão do CCL. Esse *software* emprega a morfometria baseada em *voxels* (*voxel-based morphometry*, VBM) para padronizar e simplificar as características cerebrais, o que inclui garantir uma forma cerebral uniforme (preservando a estrutura e a topologia do cérebro), converter o volume em intensidade de sinal (modulação) e ajustar o volume cerebral de acordo com o ICV e a idade.

O *software* realiza análises baseadas em ROIs predefinidas de diferentes atlas cerebrais, como o Atlas Automatizado de Rotulagem Anatômica (108 ROIs), o Atlas de Brodmann (118 ROIs) e o Atlas Probabilístico do LONI (56 ROIs), ajustando os volumes locais pelo ICV e pela idade. O uso de múltiplos atlas permite que o aprendizado de máquina aproveite a redundância e a sobreposição entre as ROIs para melhorar a precisão e eficácia na identificação de padrões cerebrais.

Para desenvolver o modelo de classificação, foi utilizado um classificador SVM treinado e validado com dados da base ADNI (359 DA, 284 CCLp (87,3% converteram em até 1 ano e 95,8% converteram em até 3 anos), 128 CCLe, que permaneceram estáveis por 4 anos ou mais, e 543 indivíduos CN). No treinamento, as imagens de DA e CCLp foram rotuladas como pertencentes ao espectro DA, enquanto as imagens de CN e CCLe foram rotuladas como pertencentes ao espectro não-DA. Foram utilizadas 321 imagens do espectro DA (179 DA e 142 CCLp) e 335 imagens do espectro não-DA (271 CN e 64 CCLe). Para comparação, foram desenvolvidos dois SVMs: um utilizando apenas estruturas cerebrais (SVMest) e outro utilizando estruturas cerebrais e informações da avaliação cognitiva MEEM (SVMcog).

Para quantificar a probabilidade de que um cérebro seja afetado pela DA, foi utilizado o *Alzheimer's Disease Likelihood Score* (ADLS), calculado a partir da função de probabilidade posterior, que estima a probabilidade de um exemplo pertencer ao espectro DA com base nas características cerebrais extraídas, sendo usado para ajustar os parâmetros do modelo. A capacidade de generalização do modelo foi verificada utilizando-se os bancos de dados *Australian Imaging, Biomarker & Lifestyle Flagship Study of Aging* (AIBL) (447 CN e 72 DA), *Japanese Alzheimer's Disease Neuroimaging Initiative* (JADNI) (165 CN, 158 DA, 129 CCLe e 140 CCLp), *Minimal Interval Resonance Imaging in Alzheimer's Disease* (MIRIAD) (23 CN e 46 DA) e *Open Access Series of Imaging Studies* (OASIS) (98 CN e 30 DA).

O modelo SVMest, para classificação em espectro DA ou não-DA utilizando as imagens da ADNI, obteve acurácia de 88,3%, sensibilidade de 81,6% e especificidade de 85,1%. Já o SVMcog apresentou acurácia de 91,9%, sensibilidade de 92,5% e especificidade de 91,4%. Quando aplicado aos demais bancos de dados, o SVMest apresentou acurácia variando de 89,2% até 97,1%, sensibilidade de 91,1% até 96,7% e especificidade de 87,3% até 100%. Enquanto o SVMcog obteve acurácia variando de 93,8% até 100%, sensibilidade de 93,3% até 100% e especificidade de 93,9% até 100%.

Para avaliar a predição da conversão do CCL, foram utilizados 142 CCLp e 64 CCLe da base de dados ADNI, considerando  $ADLS > 0,5$  como valor de corte para os SVMs. Como resultado, o SVMest obteve acurácia de 83%, sensibilidade de 81,7%, especificidade de 85,9% e F1-score de 86,9%. O SVMcog obteve acurácia de 85%, sensibilidade de 83,8%, especificidade de 87,5% e F1-score de 88,5%.

O SVMest também foi avaliado em relação ao tempo antes da conversão, mas apresentou um grande desequilíbrio de classe conforme o tempo aumentou. Como resultado, os valores de Coeficiente de Correlação de Matthews para o SVMst mostraram uma diminuição significativa entre 2 e 3 anos antes da conversão (66,7% e 53%, respectivamente), revelando uma dificuldade crescente do modelo em distinguir entre pacientes CCLe e CCLp à medida que o tempo entre a aquisição da imagem e a conversão para DA aumenta.

A definição de CCLe no estudo depende do período de acompanhamento, o que significa que alguns pacientes classificados como CCL podem eventualmente progredir para DA se forem monitorados por um período mais longo. Portanto, o modelo proposto pode não ser ideal para estratégias preventivas de longo prazo. Além disso, o fato de que as imagens cerebrais de um mesmo sujeito foram usadas tanto no conjunto de treinamento quanto no conjunto de teste pode dificultar a identificação de sobreajuste no modelo.

No estudo de Lee, Lee e Kim (2020), foi investigada a hipótese de que as texturas de imagens de RM podem prever a progressão de CCL para DA em estágio mais precoce do que os atributos volumétricos. Para isso, foram analisados o volume hipocampal, a textura das regiões hipocampais, volumes combinados e texturas combinadas, incluindo

atributos dos hipocampos, precuneus e córtex posterior cingulado. No total, 63 atributos de textura foram extraídos a partir da técnica GLCM, considerando uma distância de 1 *voxel* e 13 direções.

O conjunto de desenvolvimento incluiu 266 imagens 3T ponderadas em T1 da ADNI-2, adquiridas na linha de base ou triagem, de 145 pacientes com DA e 121 CN, acompanhados por pelo menos 2 anos. O conjunto de validação foi composto por imagens da linha de base de 194 pacientes com aCCL, dos quais 113 eram aCCLe (estáveis por pelo menos 3 anos a partir da linha de base) e 81 eram aCCLp (progrediram dentro de 3 anos). Dentro do grupo aCCLp, os pacientes foram segregadas de acordo com o tempo que levaram para converter a partir da linha de base, classificando 40 pacientes, que progrediram após 12 meses, como CCLI progressivos (CCLIp) e 41 pacientes, que progrediram dentro de 12 meses, como CCLT progressivos (CCLTp).

A seleção dos atributos foi realizada utilizando a técnica de regressão logística da biblioteca *Glmmnet*, que utiliza regressão Ridge e LASSO, considerando variáveis independentes como idade e sexo. O modelo foi treinado com validação cruzada, dividindo os dados de desenvolvimento em 10 partes, utilizando 9 para treinamento e 1 para teste. Com isso, os atributos mais relevantes foram selecionados e aplicados ao conjunto de validação, obtendo-se índice de volume hipocampal (IVH), índice de textura hipocampal (ITH), contendo 15 atributos de textura, índice de volumes combinados (IVCom) e índice de texturas combinadas (ITCom), contendo 27 atributos de textura.

As AUC das curvas ROC foram utilizadas para avaliar a precisão do modelo na previsão da conversão de CCL para a DA. Em seguida, foram aplicados os testes de DeLong para comparar essas AUCs entre os diferentes índices (IVH, IVCom, ITH e ITCom) na previsão da conversão de CCL para a DA.

Os resultados mostraram que, na discriminação entre CCLe e CCLp, CCLIp e CCLTp, o desempenho do ITCom não aumentou significativamente em relação ao ITH ( $p = 0,177$ ,  $0,272$  e  $0,338$ , respectivamente). No entanto, tanto o ITH (AUC = 0,790) quanto o ITCom (AUC = 0,811) discriminaram CCLp de CCLe melhor do que o IVH (AUC = 0,739,  $p = 0,047$  para ITH e  $p = 0,007$  para ITCom).

Porém, o ITH conseguiu prever a conversão para a DA melhor do que o IVH em CCLIp (AUC = 0,795 contra 0,726,  $p = 0,06$ ), mas não em CCLTp (AUC = 0,785 contra 0,753,  $p = 0,249$ ). O ITCom, por sua vez, foi mais preciso que o IVH tanto em CCLIp (AUC = 0,817 contra 0,726,  $p = 0,029$ ) quanto em CCLTp (AUC = 0,805 contra 0,753,  $p = 0,019$ ).

Na comparação entre IVCom e IVH, não houve aumento significativo de sensibilidade na discriminação entre CCLe e CCLp, nem entre CCLe e CCLIp ( $p = 0,136$  e  $0,589$ , respectivamente). Contudo, o ITCom demonstrou mais precisão do que o IVCom para diferenciar CCLe de CCLp (AUC = 0,811 contra 0,769,  $p = 0,06$ ) e CCLe de CCLIp (AUC = 0,817 contra 0,741,  $p = 0,027$ ). No entanto, para a distinção entre CCLe e

CCLTp, o IVCom mostrou um aumento significativo ( $p = 0,044$ ), alcançando desempenho semelhante ao ITCCom ( $AUC = 0,805$  contra  $0,797$ ,  $p = 0,709$ ).

Conclui-se, portanto, que os atributos de textura são mais sensíveis a mudanças iniciais no cérebro, sendo capazes de detectar alterações até 3 anos antes da progressão para DA, o que sugere que as mudanças de textura ocorrem antes das atrofia nas estruturas cerebrais. Além disso, as mudanças nos hipocampos ocorrem mais precocemente em comparação a outras regiões, já que os índices hipocampais apresentaram maior poder discriminativo nos estágios iniciais de CCL, enquanto os índices combinados tiveram melhor desempenho nos estágios tardios.

A Tabela 9 apresenta um sumário dos trabalhos baseados em engenharia de atributos descritos nesta subseção.

## 3.2 Trabalhos correlatos com aprendizagem profunda

Ferramentas de apoio à decisão baseadas em inteligência artificial e aprendizado de máquina têm sido validadas para uma variedade de tarefas de predição clínica. Muitos dos estudos incorporaram diversos atributos, como biomarcadores complexos e avaliações neuropsicológicas, além de utilizar uma combinação de modelos.

Bae et al. (2021) desenvolveram um modelo de aprendizagem profunda para prever a conversão de CCL para DA. Imagens de RM estrutural ponderadas em T1, obtidas da ADNI em *scanners* 1,5T ou 3T, foram usadas como entrada em uma rede neural convolucional (RNC) 3D, treinada com a técnica de aprendizado por transferência (*transfer learning*).

O modelo de *transfer learning* adotado foi o ResNet29, uma variação reduzida do ResNet50, com uma arquitetura mais enxuta e curta (cerca de 4 milhões de parâmetros treináveis) em comparação ao ResNet50 (com cerca de 23 milhões de parâmetros treináveis). O modelo gera duas pontuações de predição: a probabilidade de uma imagem ser classificada como CCLp ou CCLe.

Inicialmente, o modelo foi treinado com 2084 imagens de CN e 1406 de DA de 1080 pacientes (incluindo imagens de diferentes momentos, se disponíveis), para a tarefa fonte, aprendendo os atributos que mais distinguiram entre os grupos saudáveis e doentes. Para isso, 90% das imagens foram utilizadas para treinamento, enquanto 5% foram destinadas a teste e 5% à validação, com a divisão dos dados feita de forma aleatória. Um teste ANOVA foi realizado para verificar se havia diferenças significativas nas características demográficas e nas pontuações clínicas entre os conjuntos.

Posteriormente, o modelo foi “re-treinado” utilizando imagens de RM (apenas uma por paciente) de 228 pacientes CCLp com conversão até 3 anos e 222 CCLe com permanência em CCL por pelo menos 3 anos. O objetivo foi extrair atributos capazes de prever a

Tabela 9 – Trabalhos correlatos baseados em engenharia de atributos.

	Base de dados	Qtde. imagens		Atributos extraídos	Regiões	Algoritmo de Classificação	Resultados (%)			
		CCL	CCLP				AUC	Acur.	Espec.	Sens.
Shin et al. (2021)	ADNI	203	154	Histograma Haralick RLM GLSZM MN APOE4	GM WM CSF	SVM	80,7	N.I.	73,8	74,5
Xiao et al. (2021)	ADNI	45	51	Volume: GM	Hipocampo Amígdalas Giro PH Giro TM Precuneus Giro OS Giro pré-central Insula Angular Cúneo Outras	Regressão logística esparsa com elastic net	N.I.	75,9	77,5	73,7
Minhas et al. (2021)	ADNI-1	70 <sup>1</sup>	49 <sup>1</sup>	MN <sup>1,2</sup> Volume <sup>1</sup> Área de superfície <sup>1</sup> Espessura cortical <sup>1</sup>	Amígdalas <sup>1</sup> Entorinal <sup>1</sup> Istmo do cíngulo <sup>2</sup> Fusiforme <sup>2</sup> Outras (N.I.) <sup>1,2</sup>	SVM	91,2 <sup>1</sup>	84,9 <sup>1</sup>	87,1 <sup>1</sup>	81,7 <sup>1</sup>
Syarifullah et al. (2021)	ADNI (1, 2 e GO)	128	284	Volume <sup>3,4</sup> MN <sup>4</sup>	N.I.	SVM	90,6 <sup>3</sup> 90,2 <sup>4</sup>	83,0 <sup>3</sup> 85,0 <sup>4</sup>	85,9 <sup>3</sup> 87,5 <sup>4</sup>	81,7 <sup>3</sup> 83,8 <sup>4</sup>
Calandrelli et al. (2022)	MC-PUAG	54	79	Volume: hipocampo <sup>5</sup> Espessura: córtex <sup>5</sup>	LTM TA FI Hipocampo	N.A.	77,0 <sup>5</sup>	N.I.	67,0 <sup>5</sup>	76,0 <sup>5</sup>
Xie et al. (2023)	ADNI (2 e GO)	168 <sup>6</sup>	191 <sup>7</sup>	Volumes: hipocampo anterior/posterior, PHC e ICV Espessura: CER, AB35 e AB36	Hipocampo PHC CER AB35 AB36	Modelos lineares mistos combinados com método <i>stepwise</i>	88,0 <sup>6</sup>	N.I.	N.I.	N.I.
Lee et al. (2020)	ADNI-2	113	81	Volume Textura Idade Sexo	Hipocampo Precuneus CPC	Regressão logística com Ridge e LASSO	79,0 <sup>8</sup> 81,1 <sup>9</sup> 73,9 <sup>10</sup> 76,9 <sup>11</sup>	N.I.	N.I.	N.I.

N.I.: Não informado; <sup>1</sup> valor para predição com 6 meses de antecedência (modelo NMI+RMJ); <sup>2</sup> valor para predição com 1 ano de antecedência (modelo NM); <sup>3</sup> valores para SVMest; <sup>4</sup> valores para SVMlog; <sup>5</sup> Valor referente ao melhor resultado usando variáveis quantitativas; <sup>6</sup> CCL beta-amiloide (+); <sup>7</sup> CCL beta-amiloide (-); PH: Parahipocampal; TM: Temporal Médio; OS: Occipital Superior; CPC: Córtex Posterior Cingulado; <sup>8</sup> LTH; <sup>9</sup> ITCom; <sup>10</sup> IVH; <sup>11</sup> IVCCom

conversão para DA na tarefa alvo. Para essa fase, os dados foram divididos em 70% para treinamento, 15% para validação e 15% para teste. Novamente, foi realizado um teste ANOVA para verificar diferenças significativas entre os grupos nos conjuntos obtidos. No treinamento da tarefa alvo, 127 das 155 camadas foram congeladas, resultando em 2.767.106 parâmetros treináveis. O modelo obteve uma acurácia de 82,4% e AUC de 82,7%.

Além disso, o estudo identificou, por meio de um mapa de oclusão aplicado às imagens do conjunto de teste da tarefa alvo, que regiões como o mesencéfalo, núcleo accumbens, núcleo caudado, cerebelo, globo pálido e tálamo contribuíram significativamente para a predição de CCLp, enquanto a ponte, amígdala, hipocampo e giro parahipocampal foram importantes para a predição de CCLe. Essas regiões apresentaram correlação significativa com as taxas de mudança nas avaliações clínicas e com biomarcadores do CSF.

Os autores também avaliaram o desempenho do modelo para diferentes tempos de conversão (até 3 anos, entre 3 a 6 anos e entre 6 e 10 anos) e observaram sensibilidade de 81,08%, 71,79% e 55,56%, respectivamente. Isso indica que a pontuação de predição diminui com o aumento dos tempos de conversão, gerando mais falsos negativos e, consequentemente, uma queda do desempenho do modelo. Além disso, foi observada uma diferença significativa no desempenho do modelo ao comparar a classificação de imagens 1,5T e 3T (do conjunto de teste), obtendo-se acurácia de 78,13% e 86,11%, respectivamente.

O estudo de Kung et al. (2021) utilizou atributos de volume, área de superfície e curvatura de superfície de 12 subcampos hipocampais (presubículo, subículo, CA1, CA3 e CA4 dos hemisférios esquerdo e direito), da fissura e cauda hipocampais, da camada molecular do giro denteado (GC-ML-DG), além da camada molecular do giro denteado, para modelar variações na superfície cortical e identificar a conversão de CCL para DA.

Para quantificar os padrões de dobramento dos giros e sulcos corticais, foram analisadas as curvaturas máxima e mínima (principais curvaturas) em pontos específicos da superfície de cada região. Além disso, o estudo propôs um novo biomarcador, a Razão das Curvaturas Principais (*Ratio of Principal Curvatures*, RPC), que auxilia na distinção entre áreas de dobramento e achatamento. Como atributos finais, os autores utilizaram a taxa de mudança de cada índice morfológico ao longo do tempo.

Foram usadas imagens de RM ponderadas em T1 de *scanners* de 1.5T e 3T. Os pacientes que desenvolveram DA dentro de 2 anos foram considerados CCLp. Dois métodos distintos foram empregados para a seleção dos atributos: seleção univariada, com o teste *t* de *Student* para selecionar os índices de mudanças mais divergentes entre os grupos CCLp e CCLe, e a importância dos atributos, que envolveu o treinamento de 100 florestas aleatórias com 20 árvores de decisão cada, baseadas na combinabilidade do conjunto de atributos e na impureza de Gini.

Os subcampos hipocampais identificados e seus índices estruturais, sensíveis às mu-

danças associadas à progressão de CCL para DA, foram posteriormente alimentados em um classificador perceptron multicamada (PMC) para diferenciar CCLp de CCLe. O PMC, uma rede multicamadas, permite a extração de informações em múltiplos níveis por meio de sua estrutura hierárquica. O PMC proposto pelo estudo consistia em duas camadas ocultas, cada uma com 12 neurônios.

Os autores avaliaram o desempenho do classificador PMC considerando diferentes conjuntos de atributos obtidos por meio de distintos critérios de seleção, além de avaliar o desempenho sem qualquer processo de seleção de atributos. Entre os critérios aplicados na seleção univariada, o valor- $p < 1$ , que incluiu 19 atributos, apresentou o melhor desempenho do modelo, com uma acurácia de 79,07%, sensibilidade de 72,41% e especificidade de 85,74%, superando os resultados obtidos com valores- $p$  menores que 0,05 e 0,01. No caso dos métodos baseados na importância dos atributos, a floresta aleatória que identificou os 20 atributos mais relevantes proporcionou o melhor desempenho ao PMC, com acurácia de 79,95%, sensibilidade de 74,44% e especificidade de 85,46%. Esse desempenho foi superior ao alcançado ao identificar 10, 15 e 25 atributos. O pior desempenho do modelo foi observado ao utilizar os 45 atributos (sem seleção), com uma acurácia de 65,97%, sensibilidade de 60,56% e especificidade de 71,39%.

Os autores também identificaram uma alta correlação entre os biomarcadores extraídos e as pontuações do teste neuropsicológico MEEM. Atributos de volume e área de superfície em quatro regiões (presúbulo, subúbulo, camada molecular e fissura hipocampal) mostraram correlações fortes com MEEM, enquanto a RPC foi significativa para duas regiões (CA1 e camada molecular), e as curvaturas principais foram significativas para a região CA1.

Já o estudo de Ashtari-Majlan, Seifi e Dehshibi (2022) propôs uma RNC profunda de múltiplos fluxos, alimentada com dados de imagem baseados em *patches* 3D para classificar CCLe e CCLp. Para isso, foram utilizadas imagens ponderadas em T1 de 695 pacientes (231 CN, 100 CCLe, 164 CCLp e 200 DA) da ADNI-1.

Semelhante ao trabalho de Bae et al. (2021), inicialmente Ashtari-Majlan, Seifi e Dehshibi (2022) compararam imagens DA e CN para identificar regiões cerebrais com diferenças anatômicas. Para isso, os pacientes CN e DA foram aleatoriamente divididos em conjuntos de treino (70%), validação (10%) e teste (20%). O método de detecção de marcadores baseou-se em *patches* de  $5 \times 5 \times 5$ , com cada um representado por um vetor com 29 dimensões, incluindo atributos de GLCM, Índice de Similaridade Estrutural, Erro Quadrático Médio, entropia, além da média e desvio padrão dos *voxels* do *patch*.

Em seguida, o teste estatístico T2 de Hotelling foi aplicado para gerar um mapa de p-valores, e marcadores com distância Euclidiana menor que 15 foram excluídos para evitar redundâncias. Os 50 marcadores com os menores p-valores foram selecionados e, em seguida, o tamanho dos *patches* foi ajustado para  $19 \times 19 \times 19$  para compensar áreas descartadas na detecção de marcos anatômicos.

Esses *patches* alimentaram a RNC profunda de 50 fluxos, cada fluxo composto por cinco camadas convolucionais seguidas de funções de ativação, camadas totalmente conectadas e uma camada de *dropout*. As saídas dos fluxos foram concatenadas e passaram por camadas de *dropout*, camadas totalmente conectadas e, por fim, pela função *softmax* para a classificação final. *Patches* foram selecionados aleatoriamente para alimentar os fluxos.

Após ser treinado usando *patches* de imagens DA e CN usando validação cruzada de 5 partições, os pesos aprendidos foram transferidos para a arquitetura proposta, congelando-se as camadas iniciais e treinando apenas as últimas três camadas totalmente conectadas com *patches* das imagens de pacientes CCLe e CCLp.

O modelo alcançou uma acurácia de 79,90%, sensibilidade de 75,55%, especificidade de 99,70%, F1-score de 85,96% e AUC de 94,39% na tarefa de classificação de CCLe e CCLp.

O trabalho de Luo et al. (2023) propôs um novo método de transferência de atenção, chamado “Transferência de Atenção por Ativação de Classe” (*Class Activation Attention Transfer*, CAAT), para prever a progressão de pacientes CCL para DA em um período de 3 anos, utilizando uma RNC 3D.

Inicialmente, o modelo foi treinado para aprender automaticamente ROIs relevantes a partir de três tarefas fontes de classificação: DA *vs* CN, usando 508 imagens de cada classe; pontuações altas *vs* baixas de ADAS-cog; e pontuações altas *vs* baixas de CDR-SB. Para essas últimas tarefas, foram utilizadas 1.243 imagens (382 de CCL, 460 de CN e 401 de DA) no treinamento e 310 imagens (93 de CCL, 109 de CN e 108 de DA) para o teste. Como saída dessas tarefas, foram gerados Mapas de Ativação de Classe (MACs).

Para isso, foi utilizado um modelo composto por 5 camadas convolucionais, sendo a última seguida por uma camada de *Global Average Pooling*, que calculava as médias dos mapas de características obtidos da última camada convolucional, gerando um vetor que foi ponderado pelos pesos aprendidos para prever cada classe. A partir disso, os MACs foram gerados pela ponderação de cada mapa de característica da última camada convolucional com seus pesos correspondentes. Após redimensionamento, esses MACs foram sobrepostos às imagens originais, criando mapas de calor que indicavam as ROIs aprendidas.

Após treinar o modelo para as tarefas fontes, 298 imagens de CCLe e 295 de CCLp foram usadas como entrada para extração dos MACs correspondentes. Depois disso, um segundo modelo foi treinado para classificação de CCLe e CCLp utilizando imagens dessas classes. As imagens passaram por quatro camadas convolucionais, e da última camada foram extraídos mapas de características, cujas dimensões serviram de referência para o redimensionamento dos MACs extraídos pelo primeiro modelo (das tarefas fontes). Após o ajuste de dimensões, os mapas de características e os MACs foram sobrepostos, criando uma versão ajustada que indicava as ROIs mais discriminativas. Tanto os mapas

de características originais (da 4ª camada convolucional) quanto os ajustados foram então encaminhados para uma quinta camada convolucional, seguida por duas camadas totalmente conectadas, gerando a saída de classificação (CCLe ou CCLp).

Como resultado da classificação de CCLp e CCLe, o modelo alcançou 74,61% de acurácia, 78% de AUC, 75% de sensibilidade, 75% de especificidade e 75% de F1-score.

Gao et al. (2020) também propuseram um modelo de transferência de aprendizado para prever a conversão de CCL para DA em um período de 3 anos, utilizando um biomarcador relacionado à idade. Para isso, desenvolveram uma rede neural ajustada para idade (*Age-adjust neural network*, AD-NET), baseada na hipótese de que o cérebro de um paciente CCL aparenta ser mais envelhecido em comparação ao de um paciente CN da mesma idade.

O desenvolvimento da AD-NET começou com o treinamento de uma RNC utilizando 763 imagens de RM de pacientes CN (com 84 imagens CN reservadas para teste), provenientes das bases de dados ADNI e IXI, para a extração de atributos e predição da idade.

Em seguida, para a predição de CCLp vs CCLe, 297 imagens CCL (168 CCLp e 129 CCLe) da ADNI foram utilizadas como entrada no modelo de predição da idade. A diferença entre a idade prevista e a idade real do paciente, chamada de  $\Delta_{idade}$ , foi calculada para cada indivíduo. Os autores postularam que um maior valor de  $\Delta_{idade}$  indicaria uma maior probabilidade de o paciente com CCL converter para DA.

Subsequentemente, foi criado um segundo modelo, ajustado a partir do modelo pré-treinado, com o objetivo de prever a conversão de CCL para DA (tarefa alvo). Os parâmetros do modelo pré-treinado foram mantidos, e uma camada adicional foi incorporada e ajustada com base no  $\Delta_{idade}$ , a fim de adaptar o modelo à nova tarefa de previsão da conversão de CCL. Para a classificação de CCLe e CCLp, o modelo alcançou 81% de AUC, 76% de acurácia, 77% de sensibilidade e 76% de especificidade.

Já o trabalho de Rahim et al. (2025) propõe uma arquitetura baseada em aprendizado profundo para prever a conversão de comprometimento CCLp para DA, a partir de imagens de RM. A proposta explora o potencial diagnóstico de modelos de ensemble homogêneos e heterogêneos, combinando diferentes arquiteturas de redes neurais profundas para aprimorar a acurácia na detecção da progressão da doença. Foram conduzidos experimentos utilizando planos de imagem individuais (axial, coronal e sagital) a fim de avaliar o impacto de cada orientação anatômica no desempenho do modelo.

Cinco arquiteturas convolucionais profundas — EfficientNet, XceptionNet, ResNet, ConvNext e DenseNet — foram empregadas como extratoras de atributos, sendo combinadas com três cabeçalhos classificadores otimizados por busca bayesiana: PMC, *Long Short-Term Memory* (LSTM) e Atenção Multi-Cabeça (MHA). As melhores combinações entre extratores e classificadores foram integradas em modelos de ensemble, com o objetivo de avaliar o poder coletivo de decisão das redes.

A arquitetura proposta inclui uma etapa de refinamento dos atributos extraídos por meio de um módulo de atenção *Convolutional Block Attention Module* (CBAM), seguido por classificadores que capturam dependências espaciais e temporais entre os dados. As RNCs extraíram representações hierárquicas complexas a partir das imagens, enquanto o PMC processou essas representações de forma densa, o LSTM modelou padrões sequenciais e a MHA capturou relações de longa distância entre os atributos.

O estudo utilizou dados de 192 participantes do banco ADNI, sendo 96 pacientes com CCLe e 96 com CCLp ou DA, conforme a evolução diagnóstica ao longo de 48 meses. O grupo CCLp incluiu indivíduos que converteram para DA entre os meses 18 e 48, complementado por pacientes já diagnosticados com DA para fortalecer a aprendizagem dos padrões de atrofia. As imagens utilizadas foram sequências anatômicas T1 (3T), com 132 fatias centrais extraídas dos planos axial, coronal e sagital, por conterem regiões cerebrais críticas, como hipocampo, amígdala e ventrículos.

A arquitetura proposta seleciona, para cada plano de visualização, uma combinação específica de extrator de características (uma RNC) e cabeçalho classificador (PMC, LSTM ou MHA). Devido à complexidade da arquitetura de ensemble, a análise de interpretabilidade foi concentrada nos modelos de melhor desempenho identificados: EfficientNet + MHA para o plano sagital, ConvNext + LSTM para o plano coronal e DenseNet + LSTM para o plano axial.

Os mapas de atenção gerados por esses modelos revelaram que, em pacientes CCLe, as áreas mais relevantes incluíram o hipocampo rostral, amígdala, globo pálido, amígdala lateral, córtex entorrinal e o giro parahipocampal caudal, que se mantêm relativamente preservadas nesses casos. Por outro lado, em pacientes CCLp/DA, o modelo identificou atrofia mais acentuada em regiões como o hipocampo caudal, amígdala medial, putâmen dorsolateral, giro fusiforme, globo pálido, giro parahipocampal posterior, sulco temporal superior, giro temporal médio e inferior.

Entre todos os modelos avaliados, o ensemble heterogêneo, que combina EfficientNet, XceptionNet e DenseNet com MHA como cabeçalho de classificação, obteve o melhor desempenho. Esse modelo alcançou uma acurácia de 91,77%, sensibilidade de 87,69%, especificidade de 88,53% e AUC de 92,87% no conjunto de dados da ADNI.

A Tabela 10 apresenta um sumário dos trabalhos baseados em aprendizagem profunda descritos nesta subseção.

### 3.2.1 Discussão e Considerações Finais

Este capítulo apresentou uma revisão de diferentes trabalhos propostos na literatura nos últimos anos para a extração de atributos relevantes de imagens de RM, com o objetivo de auxiliar no diagnóstico da DA. A partir dos trabalhos apresentados, pode-se destacar a limitada quantidade de imagens disponíveis para esse tipo de pesquisa e o curto período

Tabela 10 – Trabalhos correlatos baseados em aprendizagem profunda.

	Base de dados	Qtde. imagens		Atributos	Modelo	Resultados (%)			
		CCLe	CCLp			AUC	Acur.	Espec.	Sens.
Bae et al. (2021)	ADNI	222	228	Imagens inteiras de RM de: DA, CN, CCLe e CCLp	RNC com ResNet29	82,7	82,4	N.I.	N.I.
Kung et al. (2021)	ADNI (2 e GO)	89	89	Volume, área de superfície e curvatura de superfície de 12 subcampos hipocampus	PMC	N.I.	79,95 <sup>1</sup>	85,46 <sup>1</sup>	74,44 <sup>1</sup>
Ashtari-Majlan, Seifi e Dehshibi (2022)	ADNI 1	100	164	GLCM, ISE, EQM, entropia, média e desvio padrão de <i>patches</i> 3D de CN, DA, CCLe e CCLp	RNC de 50 fluxos	94,39	79,90	99,70	75,55
Luo et al. (2023)	ADNI (1 e 2)	298	295	Imagens inteiras de RM (de DA, CN, CCLe e CCLp) Pontuações neuropsicológicas	RNC com CAAT	78	74,61	75	75
Gao et al. (2020)	ADNI IXI	129	168	Idade e imagens de RM de CN, CCLe e CCLp.	AD-NET	81	76	76	77
Rahim et al. (2025)	ADNI	96	96	Imagens 2D de cada plano de visualização (axial, coronal e sagital).	EfficientNet, XceptionNet e DenseNet com MHA	88,53	91,77	88,53	87,69

N.I.: Não informado; PMC: classificador Perceptron Multicamada; <sup>1</sup> valor para seleção de atributos a partir de floresta aleatória que identificou 20 atributos mais importantes; ISE: Índice de Similaridade Estrutural; EQM: Erro Quadrático Médio; CAAT: Transferência de Atenção por Ativação de Classe.

Fonte: Autoria própria.

de acompanhamento dos pacientes, o que constitui uma limitação para a generalização dos modelos propostos.

Além disso, observa-se que as regiões do LTM, como o hipocampo, a amígdala, o giro parahipocampal e a região entorrinal, são as mais exploradas (MINHAS et al., 2021; XIAO et al., 2021; CALANDRELLI et al., 2022; XIE et al., 2023), evidenciando a sua possível relação com a progressão da DA. Outro ponto relevante é a variabilidade dos atributos estruturais utilizados entre os estudos, incluindo volumes, espessuras, áreas de superfície, distribuição dos valores dos *pixels* e textura, sendo os dois primeiros os mais frequentes nos trabalhos revisados (XIAO et al., 2021; MINHAS et al., 2021; SYAIFULLAH et al., 2021; CALANDRELLI et al., 2022; XIE et al., 2023).

A combinação de regiões e atributos extraídos variou consideravelmente entre os estudos sobre a progressão da DA, refletindo uma ampla diversidade nas abordagens metodológicas. Enquanto alguns estudos se concentraram exclusivamente em medidas de GM, outros incluíram dados de todos os três tipos de tecido cerebral (GM, WM e CSF). Além disso, houve uma variação significativa na escolha das ROIs: alguns estudos utilizaram um número limitado de regiões, enquanto outros aplicaram múltiplas ROIs de atlas sem especificar quais regiões foram relevantes.

Os modelos propostos baseados em *deep learning*, apesar de terem apresentado resultados comparáveis aos trabalhos que utilizaram engenharia de atributos, apresentaram baixa interpretabilidade, não esclarecendo quais os atributos anatômicos e ROIs relevantes para compreender a progressão da DA.

Embora alguns estudos tenham abordado medidas específicas como volume, espessura, textura e área de superfície, poucos investigaram todas essas medidas de forma abrangente em diversas ROIs e para todos os tecidos. Essa diversidade metodológica e a ausência de uma abordagem uniforme nas análises representam uma limitação significativa para a compreensão de quais atributos e regiões estão mais fortemente associados à progressão de CCL para DA.

No Capítulo 4 serão apresentados os materiais e métodos empregados para a realização desta pesquisa.



---

# Capítulo 4

## Materiais e Métodos

---

*Este capítulo apresenta uma descrição das bases de dados que foram utilizadas, bem como das técnicas que foram empregadas neste trabalho, visando facilitar a compreensão deste trabalho.*

### 4.1 Visão Geral

Neste projeto de pesquisa, foi desenvolvido um modelo de Floresta Aleatória baseado em engenharia de atributos, com o objetivo de classificar casos de CCL em CCLe e CCLp, a partir de imagens de RM. O modelo foi construído utilizando atributos derivados da espessura cortical (total e por ROI), volume de cada ROI (total e por tipo de tecido), além de medidas de textura extraídas com base na GLCM. O fluxograma apresentado na Figura 18 resume as principais etapas realizadas.

### 4.2 Base de dados

Para a realização desta pesquisa foram utilizadas 2 bases de dados contendo imagens 3D de RM ponderadas em T1. Uma delas fornece modelos geométricos (malhas) de estruturas cerebrais importantes, como os hipocampus e amígdalas, enquanto a outra possui imagens de pacientes diagnosticados com CCLp e CCLe. As bases de dados são descritas em detalhes a seguir.

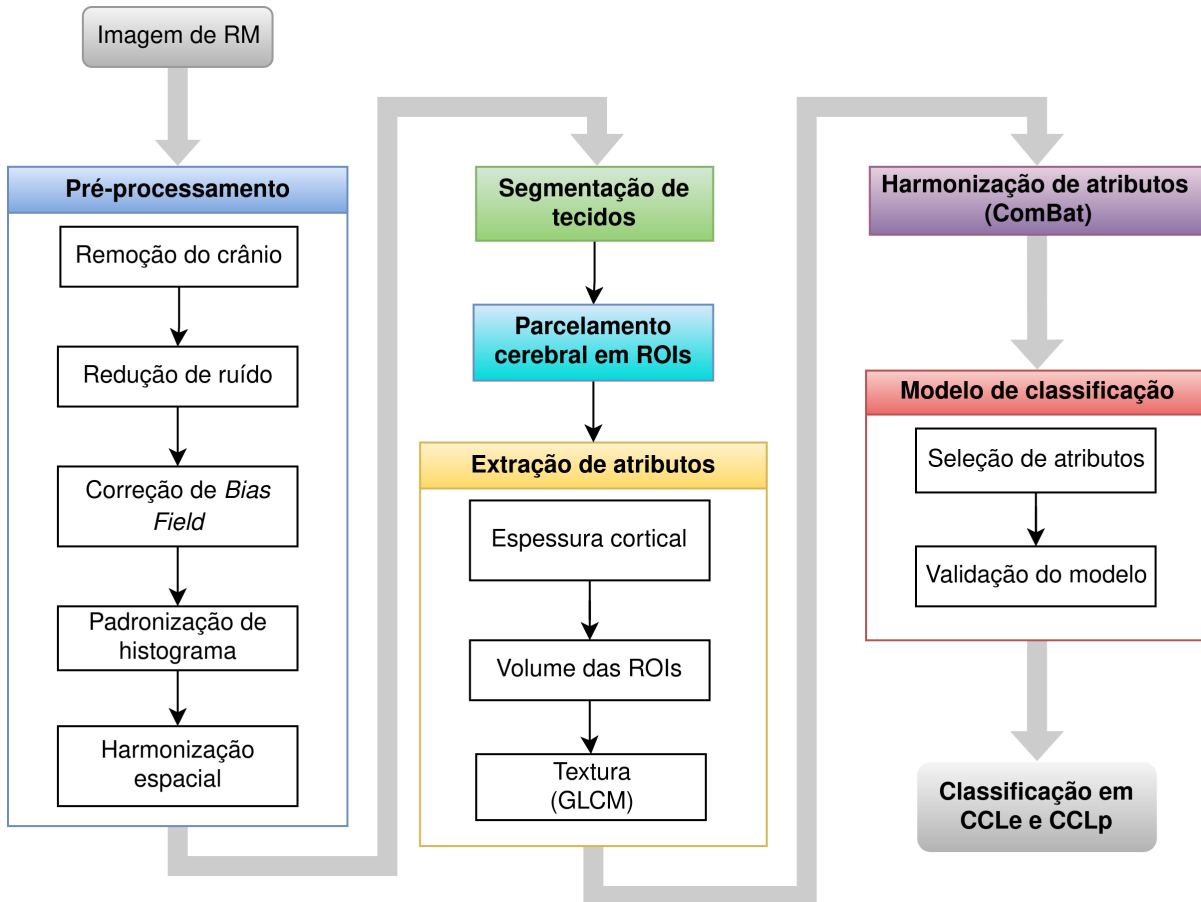
#### 4.2.1 ADNI

Parte dos dados utilizados na preparação deste trabalho foram obtidos do banco de dados da *Alzheimer's Disease Neuroimaging Initiative* (ADNI)<sup>1</sup>. A ADNI foi lançada em

---

<sup>1</sup> <http://adni.loni.usc.edu/>

Figura 18 – Fluxograma das principais etapas realizadas neste projeto de pesquisa.



Fonte: Autoria própria.

2003 como uma parceria público-privada, liderada pelo investigador principal Michael W. Weiner, MD. O objetivo principal da ADNI tem sido testar se imagens seriadas de RM, PET, outros marcadores biológicos e avaliação clínica e neuropsicológica podem ser combinados para medir a DA precoce e a progressão do CCL.

Os estudos da ADNI são realizados por pesquisadores-colaboradores e utilizam centros norte-americanos para coleta das informações dos pacientes, como idade, sexo, tempo de escolaridade, classificação do estágio cognitivo do paciente, informações genéticas, resultados de exames neuropsicológicos, imagens de RM, ressonância magnética funcional, PET e Imagem por Tensor de Difusão com ponderações em T1, T2 e DP.

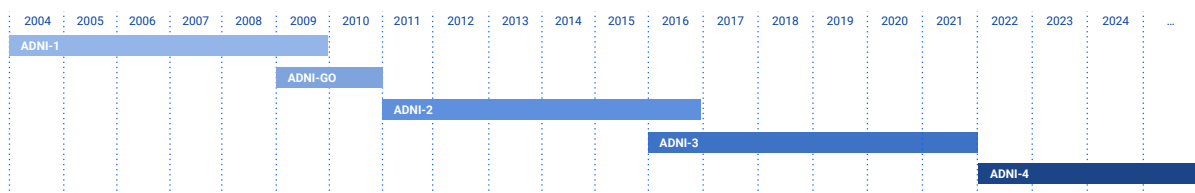
Essas informações compõem o banco de dados, disponibilizado gratuitamente pelo Laboratório de Neuroimagem (*Laboratory of Neuro Imaging, LONI*) da Universidade do Sul da Califórnia (*University of Southern California, USC*) na página de *Image & Data Archive (IDA)*<sup>2</sup>.

Desde a sua fundação, já foram concluídas quatro fases de estudos (conforme Figura 19), construindo uma base de dados cada vez mais informativa e representativa. Segundo a

<sup>2</sup> <https://ida.loni.usc.edu/login.jsp?project=ADNI>

ADNI (2024), a primeira fase (ADNI-1) iniciou-se em 2004 e contou com 400 participantes com CCL, 200 com DA e 200 CN, todos com idade entre 55 e 90 anos. A segunda fase (ADNI-GO), iniciada em 2009, adicionou 200 participantes com CCLI à ADNI-1 a fim de examinar os biomarcadores em um estágio anterior. Já em 2011, iniciou-se a terceira fase (ADNI-2), que adicionou à ADNI-1 mais 150 CN, 100 com CCLI, 150 com CCLT e 150 pacientes com DA leve. Em seguida, em 2016, iniciou-se a quarta fase (ADNI-3), na qual adicionou-se 133 CN, 151 com CCL e 87 pacientes com DA. Atualmente, encontra-se em andamento a quinta fase (ADNI-4), iniciada em 2022, cujo objetivo principal é acrescentar 750 novos pacientes e incluir 750 pacientes remanescentes do ADNI-3 visando a inclusão de populações sub-representadas para melhorar a generalização dos resultados e o entendimento da disparidade da saúde entre as populações. Para esta última fase, o tempo de estudo previsto é de cinco anos.

Figura 19 – Evolução longitudinal da base de dados da ADNI.



Fonte: Informações extraídas de ADNI (2024).

Para este estudo, foram utilizadas imagens de RM estrutural ponderadas em T1, obtidas pela sequência *Magnetization Prepared Rapid Gradient Echo* (MPRAGE) a partir de *scanners* de 1,5T e 3T. As imagens da sequência MPRAGE foram preferidas por seguirem as recomendações da ADNI, que afirma que os arquivos MPRAGE são os melhores em termos de qualidade de imagem, pois passaram por *gradwarping* (correção de distorções geométricas causadas por gradientes de campo magnético não uniformes) e correção de intensidade, além de serem escalados para o desvio de gradiente utilizando os dados do fantoma.

As imagens da ADNI podem apresentar cinco diagnósticos distintos: Preocupação Significativa com a Memória (PSM, *Significant Memory Concern*), CN, CCLI, CCLT e DA. Para a distinção entre CCLI e CCLT, a ADNI utiliza a Escala de Memória de Wechsler (teste de avaliação psicológica que mede diferentes aspectos da memória, como memória verbal e visual, reconhecimento de figuras, entre outros) em conjunto com a subescala Memória Lógica II, que avalia a capacidade de lembrar e reproduzir um texto narrativo após um período de atraso, conforme Tabela 11. Para os fins deste estudo, optou-se por unificar as categorias CCLI e CCLT na categoria CCL, resultando em 4 grupos diagnósticos para as imagens: PSM, CN, CCL e DA.

Para compor o grupo CCL, para a realização deste trabalho, selecionou-se os pacientes que receberam o primeiro diagnóstico de CCL e manteve esse diagnóstico por 36 meses

Tabela 11 – Pontuações da Escala de Memória de Wechsler utilizadas pela ADNI para classificação dos pacientes em CCLI e CCLT.

Tempo de escolaridade (anos)	CCLI	CCLT
0 - 7	3 - 6	$\leq 2$
8 - 15	5 - 9	$\leq 4$
$\geq 16$	9 - 11	$\leq 8$

Fonte: Informações extraídas de ADNI (2024).

(ou mais). Já para compor o grupo CCLp, selecionou-se os pacientes que receberam diagnóstico de CCL e que foi modificado para DA em até 36 meses, mantendo somente os pacientes que possuem alguma imagem CCL disponível.

Diante disso, para a execução deste trabalho, foram selecionadas 528 imagens de RM estrutural de 486 indivíduos na faixa etária entre 55 e 93 anos, divididas em dois grupos diagnósticos: CCLp e CCLe, cada um com apenas uma imagem de cada paciente, e com 42 pacientes sendo incluídos em ambos os grupos. A resolução das imagens selecionadas é de aproximadamente  $170 \text{ mm} \times 256 \text{ mm} \times 256 \text{ mm}$ , com pequenas variações nas dimensões. Os tamanhos dos *voxels* são em torno de  $1,2 \text{ mm} \times 1 \text{ mm} \times 1 \text{ mm}$  e também apresentam pequenas variações entre as imagens.

O número de pacientes, idade e sexo de cada grupo diagnóstico do conjunto de imagens utilizado são apresentados na Tabela 12 e os resultados da estratificação por faixa etária e sexo, por grupo CCLp e CCLe, são mostrados nas Figuras 20 e 21.

Tabela 12 – Conjunto de imagens da base de dados ADNI usadas neste trabalho.

	CCL estável	CCL progressivo
<b>Pacientes</b>	272	256
<b>Sexo (Feminino:Masculino)</b>	110:162	114:142
<b>Idade (anos)</b>	$73,05 \pm 7,29$	$74,71 \pm 7,37$

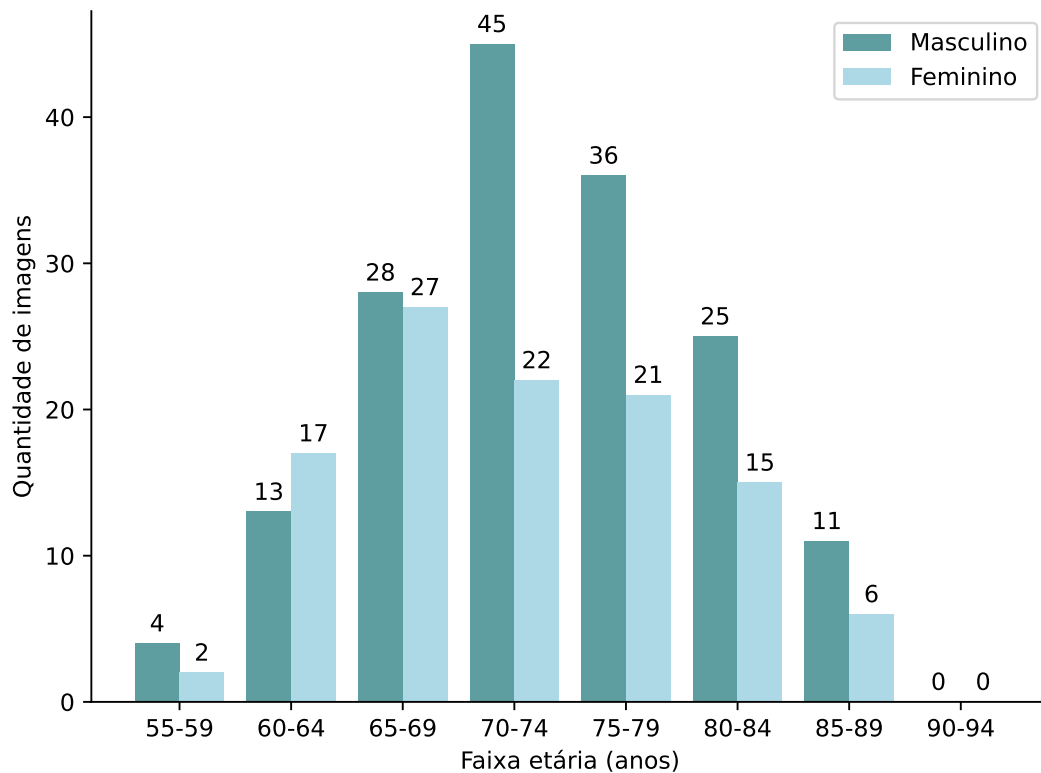
Fonte: Autoria própria.

Conforme a Figura 22, 67,97% das imagens selecionadas para o grupo CCLp foram obtidas a partir de *scanners* de 1,5T e 32,03% foram obtidas a partir de *scanners* de 3T. Para o grupo CCLe, 43,01% das imagens foram adquiridas de *scanners* de 1,5T e 56,99% foram obtidas a partir de *scanners* de 3T.

#### 4.2.1.1 Refinamento do Conjunto de Dados

Para garantir a qualidade e a consistência, todas as imagens inicialmente incluídas passaram por um processo de controle de qualidade. A extração de características foi realizada primeiramente no conjunto de dados, seguida pela aplicação do algoritmo Isolation

Figura 20 – Estratificação por idade e sexo dos pacientes CCLe da base de dados ADNI.



Fonte: Autoria própria.

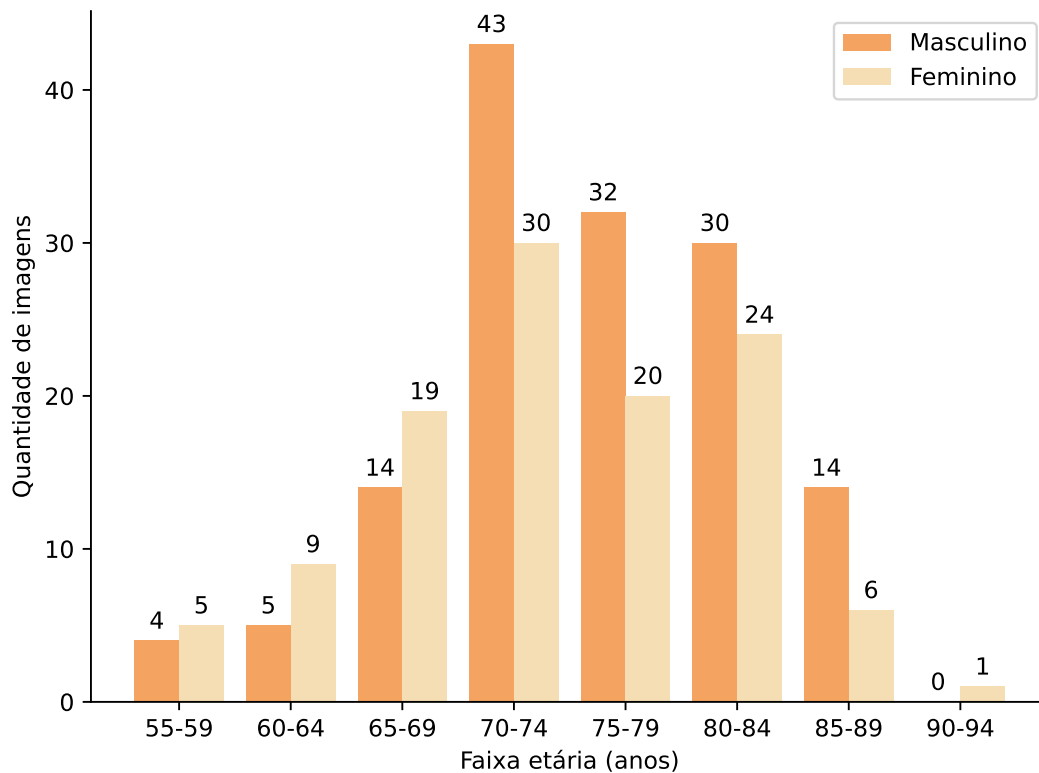
Forest para detecção de *outliers*. O modelo foi configurado com os seguintes hiperparâmetros: `contamination='auto'`, `max_samples=256` e `random_state=42`. Essa configuração permite o algoritmo estimar automaticamente a proporção de anomalias e construir um ensemble de 100 estimadores base treinados em subconjuntos aleatórios com 256 amostras cada.

Esse processo foi conduzido em três etapas: (i) nos dados originais (não harmonizados), (ii) nos dados harmonizados, (iii) nos dados harmonizados após a remoção de características que ainda apresentavam diferenças significativas relacionadas ao scanner ou ao local de aquisição.

Um total de 50 imagens foram sinalizadas como possíveis *outliers*. Cada uma delas foi então inspecionada visualmente para avaliar a integridade anatômica e a presença de artefatos. Dentre essas, 17 imagens apresentaram problemas substanciais de qualidade, como artefatos de corte ou distorções geométricas severas. Uma vez que poderiam introduzir viés nas estimativas dos modelos e afetar a robustez dos resultados, essas imagens foram removidas das análises seguintes. As 33 imagens restantes, embora inicialmente marcadas, foram julgadas como tendo qualidade aceitável e foram mantidas.

Das 17 imagens excluídas, 8 pertenciam ao grupo CCLe e 9 ao grupo CCLp, com cada exclusão correspondendo a um indivíduo distinto.

Figura 21 – Estratificação por idade e sexo dos pacientes CCLp da base de dados ADNI.



Fonte: Autoria própria.

Após esse refinamento, o conjunto final foi composto por 511 imagens de alta qualidade: 264 do grupo CCLe (51,66%) e 247 do grupo CCLp (48,34%), conforme detalhado na Tabela 13.

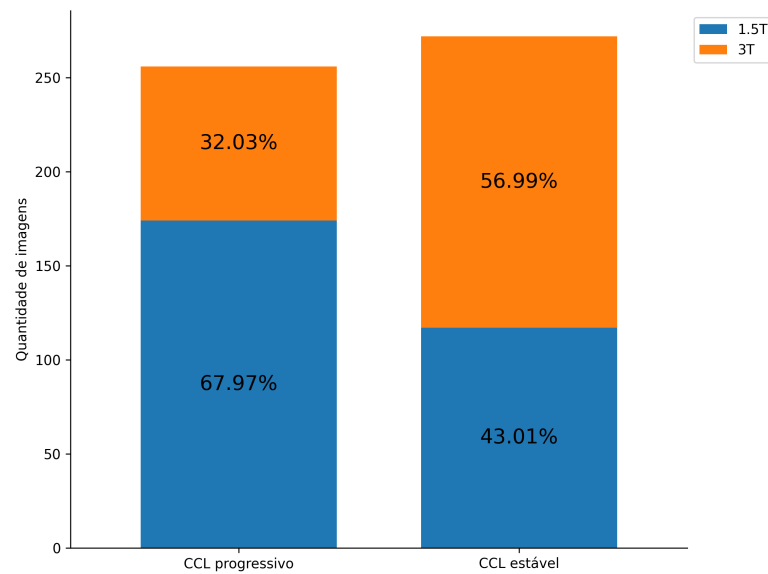
Quanto à distribuição dos *scanners*, conforme Figura 23, 67,2% das imagens do grupo CCLp foram adquiridas em *scanners* de 1,5T e 32,8% em *scanners* de 3T. No grupo CCLe, 56,4% das imagens foram obtidas com *scanners* de 3T e 43,6% com *scanners* de 1,5T.

Tabela 13 – Conjunto de imagens da base de dados ADNI usadas neste trabalho (após remoção de outliers).

	CCL estável	CCL progressivo
<b>Pacientes</b>	264	247
<b>Sexo (Feminino:Masculino)</b>	108:156	113:134
<b>Idade (anos)</b>	72,9 ± 7,26	74,59 ± 7,37

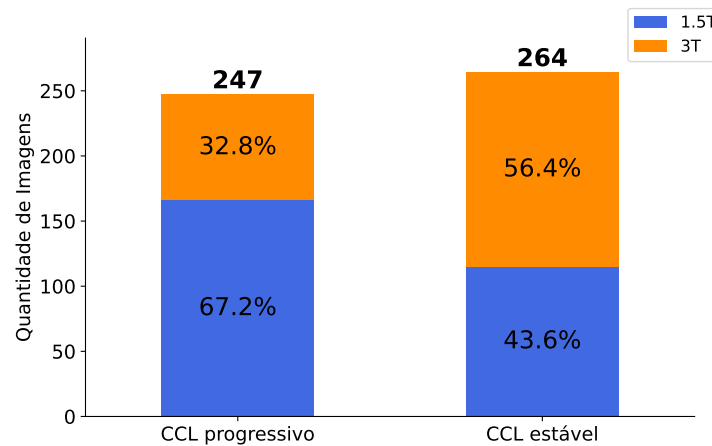
Fonte: Autoria própria.

Figura 22 – Estratificação das imagens da base de dados ADNI por intensidade de campo magnético de cada grupo diagnóstico.



Fonte: Autoria própria.

Figura 23 – Estratificação das imagens da base de dados ADNI por intensidade de campo magnético de cada grupo diagnóstico (sem *outliers*).

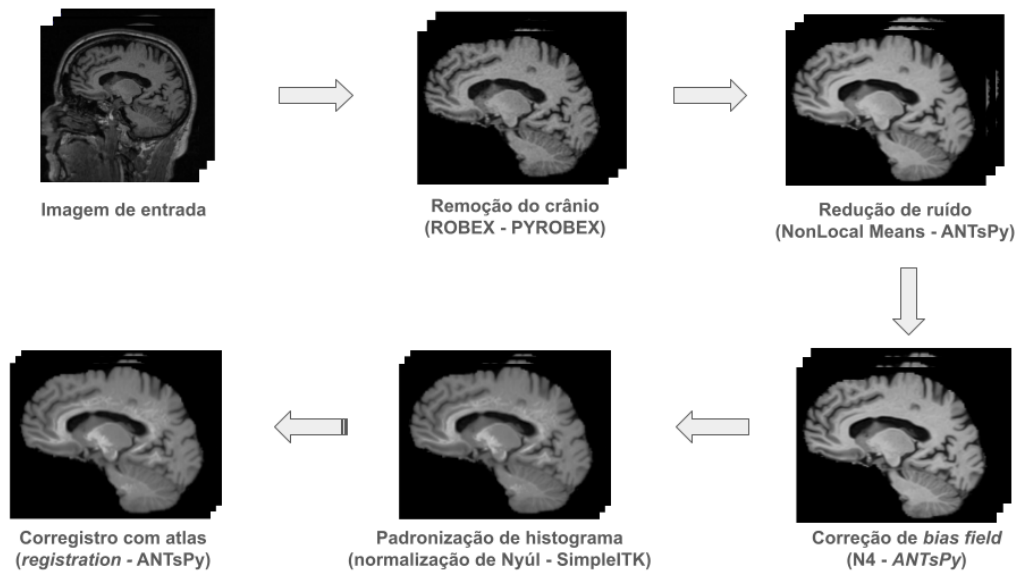


Fonte: Autoria própria.

#### 4.2.2 NAC Brain Atlas

O *Neuroimage Analysis Center* (NAC) é uma entidade de pesquisa associada à *Harvard Medical School* que disponibiliza gratuitamente uma base de dados, na qual há uma imagem de referência de RM ponderada em T1, caracterizada por uma resolução isotrópica de  $1\text{mm}^3$  e dimensões de  $256 \times 256 \times 256$  *voxels* (HALLE et al., 2017). A imagem foi obtida a partir de um voluntário masculino saudável de 42 anos em um *scanner* Siemens 3T com sequência 3D MPRAGE. Os parâmetros de aquisição incluíram: TR = 2530 ms,

Figura 24 – Etapas de pré-processamento.



Fonte: Autoria própria.

TI = 1200 ms, TE = [1,64; 3,5; 5,36; 7,22] ms, Ângulo de inclinação (*flip angle*) de 7°.

Os rótulos por *voxel* foram gerados usando a segmentação automática do Freesurfer, seguida pela segmentação manual para ajustar e refinar as regiões identificadas. Com isso, a base de dados fornece 259 modelos geométricos (malhas) de diferentes estruturas cerebrais, previamente identificados por especialistas, compondo um atlas anatómico cerebral (HALLE et al., 2017). Assim, como todas as malhas estão anatomicamente alinhadas à imagem de referência, esta foi utilizada como base para todas as etapas de pré-processamento, garantindo a padronização entre as imagens utilizadas neste estudo.

## 4.3 Pré-processamento

Antes de extrair e avaliar os atributos das imagens de RM, é essencial realizar etapas de pré-processamento para melhorar e padronizar os sinais das imagens, garantindo maior reprodutibilidade do estudo. As etapas de pré-processamento que foram aplicadas neste estudo são indicadas na Figura 24 e uma breve descrição de como cada uma foi implementada é apresentada nas subseções a seguir.

### 4.3.1 Remoção do crânio

A imagem de RM cerebral também inclui tecidos não-cerebrais, que podem introduzir ruído e comprometer a segmentação, dificultando a identificação de estruturas cerebrais e a extração de características relevantes. Para melhorar a precisão da análise das imagens de RM cerebrais, é necessário realizar a remoção do crânio. Para isso, foi utilizado o

módulo ROBEX (vide Subseção 2.2.3) do pacote PYROBEX em Python, disponível gratuitamente. Esse método foi escolhido por sua abordagem híbrida, que combina modelos discriminativos e generativos, proporcionando maior robustez e adaptabilidade ao lidar com diferentes variações de imagens de RM e melhorando a qualidade da segmentação cerebral. Essa etapa resultou tanto no volume anatômico sem o crânio quanto em uma máscara binária do encéfalo.

### 4.3.2 Redução de ruído

Como descrito na Subseção 2.2.4, os dados das imagens de RM são comumente contaminadas por “ruído branco” oriundo de fontes térmicas, como os componentes eletrônicos do equipamento de RM ou do próprio corpo do paciente posicionado no interior do *scanner* (LÓPEZ; FREDERICK; VENTURA, 2021). Isso configura um grande obstáculo para análises qualitativas e quantitativas das imagens de RM, sendo crucial realizar um eficiente processo de reconstrução das imagens utilizando métodos de redução de ruído antes das etapas de processamento.

Dessa forma, para a atenuação do ruído presente nas imagens, foi utilizada a técnica Rician-adapted NLM (BUADES; COLL; MOREL, 2005a) (vide Subsubseção 2.2.4.1), disponibilizado na biblioteca ANTsPy<sup>3</sup>. A variância do ruído foi estimada diretamente de cada imagem, e os parâmetros foram selecionados empiricamente (*patch radius* = 1 *voxel*, *search radius* = 3 *voxels*, *shrink factor* = 1) para equilibrar o desempenho da redução de ruído com a fidelidade estrutural. Essa escolha segue estudos anteriores que demonstraram que essas configurações alcançam um compromisso ideal entre a supressão de ruído e a preservação dos contornos em imagens de RM estrutural (MANJÓN et al., 2010). A estimativa automática do ruído também facilitou a adaptação aos perfis de ruído específicos de cada exame.

### 4.3.3 Correção de *bias field*

Conforme discutido na Subseção 2.2.5, as imagens de RM também são afetadas pela heterogeneidade das bobinas, variações no campo magnético estático ( $B_0$ ), entre outras variações que causam o fenômeno *bias field*. Esse fenômeno impacta negativamente a análise quantitativa das imagens de RM e, portanto, exigem uma etapa de correção.

Para isso, foi utilizado o método “*N<sub>4</sub> bias field correction*” (*N<sub>4</sub>ITK*) (TUSTISON et al., 2010) (vide Subsubseção 2.2.5.1), disponibilizado na biblioteca ANTsPy. Foi utilizado um fator de redução de 3 e 20 iterações por nível em um esquema de otimização multirresolução. A convergência foi definida por um limite de variação relativa inferior a  $1 \times 10^{-3}$ . A correção e o reescalamento de intensidade subsequente foram restritos à máscara cerebral para evitar artefatos no fundo.

<sup>3</sup> <https://antspy.readthedocs.io/en/latest/utils.html>

### 4.3.4 Padronização de histogramas

Outro desafio das imagens de RM é a falta de padronização da intensidade dos *voxels*. Devido a diversas variações relacionados ao *scanner*, os valores absolutos de intensidade dos *voxels* não são fixos, o que faz com que as imagens, até de um mesmo paciente, possam apresentar aparências distintas, dificultando as análises quantitativas.

Consequentemente, é esperado que as imagens da base de dados da ADNI apresentem uma distribuição não padronizada dos níveis de cinza para um mesmo tecido, tornando necessária a padronização dos histogramas de todas as imagens a serem utilizadas.

Para isso, a imagem da base NAC foi adotada como referência, e as demais imagens foram ajustadas utilizando o algoritmo proposto por Nyúl, Udupa e Zhang (2000) a fim de padronizar a distribuição de seus níveis de cinza. A padronização de histogramas foi aplicada utilizando o método “*Histogram Matching Image Filter*” implementado na biblioteca *Simple Insight Segmentation and Registration Toolkit* (SimpleITK). Dez pontos de referência baseados em quantis foram utilizados para alinhar os histogramas das imagens entre os pacientes. A operação foi aplicada exclusivamente às intensidades do tecido cerebral, com o objetivo de suprimir artefatos não cerebrais e garantir a consistência do contraste entre os pacientes.

Em linhas gerais, o algoritmo mencionado normaliza os valores de escala de cinza de uma imagem de RM com base nos valores de uma imagem de referência, por meio de uma técnica de correspondência de histogramas, em que os histogramas são combinados para um número especificado de valores quantis.

### 4.3.5 Harmonização espacial

A imagem de RM possui três sistemas de coordenadas: coordenada real (relativo ao mundo real, determinado por parâmetros do *scanner*), coordenada anatômica (relativo ao paciente) e coordenada da imagem. Diante disso, a origem e a orientação das imagens variam de acordo com o paciente e parâmetros do *scanner*.

Assim, para garantir um processamento correto nas etapas de processamento, faz-se necessário o alinhamento de todas as imagens em um mesmo sistema de coordenadas (origem e orientação), além de ajustá-las para as mesmas dimensões e resoluções. Para isso, as imagens pré-processadas foram submetidas a um procedimento estruturado de harmonização espacial composto por três etapas: reamostragem dos *voxels*, padronização geométrica e registro deformável entre atlas e paciente.

#### 4.3.5.1 Reamostragem dos *voxels*

Cada imagem sem o crânio e sua respectiva máscara cerebral foram reamostradas para uma resolução isotrópica de 1 mm<sup>3</sup>. Imagens com valores contínuos foram reamostradas por interpolação linear, enquanto máscaras binárias utilizaram interpolação por vizinho

mais próximo, garantindo geometria consistente dos *voxels* e preservação da integridade dos rótulos.

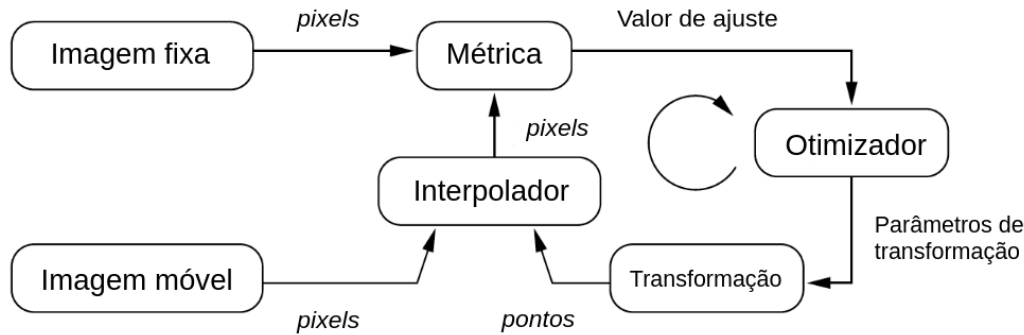
#### 4.3.5.2 Padronização geométrica

A imagem de referência NAC, sua máscara e os rótulos do atlas associados foram alinhados ao espaço nativo da imagem de cada paciente utilizando o ANTsPy2. Esse alinhamento preservou os metadados espaciais — como espaçamento entre *voxels*, cossenos diretores e origem física — sem impor orientações canônicas (por exemplo, RAS ou LPS), garantindo compatibilidade entre diferentes geometrias de aquisição. Para os pacientes que já apresentavam conformidade com a geometria de referência, foi aplicada apenas a normalização da intensidade. Máscaras de segmentação de rótulos foram excluídas desta etapa para evitar distorções causadas por interpolação.

#### 4.3.5.3 Registro deformável entre atlas e paciente

O atlas anatômico NAC foi registrado de forma não linear à imagem ponderada em T1 no espaço nativo de cada paciente, utilizando o algoritmo de *Symmetric Normalization* (SyN). O registro empregou informação mútua como métrica de similaridade e seguiu um cronograma de otimização em múltiplas resoluções com 100, 50 e 20 iterações. Além disso, optou-se por não realizar inicializações rígidas ou afins antes do SyN, a fim de evitar artefatos de interpolação e preservar a configuração espacial nativa da anatomia de cada paciente. O registro foi restrito à região intracraniana por meio da máscara cerebral da imagem fixa. Os campos de deformação resultantes foram então utilizados para deformar todos os mapas de rótulos das regiões de interesse (ROIs) definidas pelo atlas para o espaço nativo de cada paciente, utilizando interpolação por vizinho mais próximo. Isso assegurou o alinhamento espacial de regiões anatomicamente relevantes, preservando a natureza discreta dos limites das ROIs — possibilitando um parcelamento preciso e individualizado para a extração de características subsequente.

A técnica de corregristo de imagens é um processo que determina uma transformação espacial  $T$ , mapeando pontos de uma imagem para pontos homólogos em um objeto na segunda imagem, alinhando espacialmente uma imagem móvel a uma imagem fixa (JOHNSON; MCCORMICK; IBÁÑEZ, 2021). Na Figura 25, é apresentado um típico *framework* do processo de registro de imagens. O componente de transformação  $T$  mapeia pontos da imagem fixa para a imagem móvel. O interpolador avalia as intensidades da imagem móvel em posições fora da grade, enquanto o componente métrico avalia a correspondência entre a imagem fixa e a imagem móvel transformada, servindo como critério quantitativo para otimização, baseado nos parâmetros da transformação (JOHNSON; MCCORMICK; IBÁÑEZ, 2021).

Figura 25 – *Framework* de registro de imagens.

Fonte: Imagem extraída e modificada de Johnson, McCormick e Ibáñez (2021).

## 4.4 Segmentação de tecidos

Para diferenciar os três tecidos cerebrais (GM, WM e CSF) em cada uma das imagens de RM, foi utilizado o algoritmo Atropos, que faz parte do pacote *Advanced Normalization Tools* (ANTs) do Python e que utiliza o ITK como base de desenvolvimento.

O Atropos é um algoritmo de segmentação automática de tecidos composto por três componentes principais: modelos de verossimilhança, probabilidade *a priori* derivadas de um Campo Aleatório de Markov (*Markov Random Field*, MRF), e um *framework* de otimização para maximizar a probabilidade *a posteriori* (AVANTS et al., 2011).

Inicialmente, as intensidades dos *voxels* da imagem de entrada são modeladas por um GMM, que assume que os dados são explicados por uma combinação de várias distribuições Gaussianas, cada uma com seus próprios parâmetros. O objetivo é ajustar esses parâmetros (médias, variâncias e pesos) de cada componente do GMM de modo a maximizar a probabilidade de que cada intensidade de *voxel* pertença a uma das classes, com o peso atribuído refletindo a proporção de cada componente (AVANTS et al., 2011).

Entretanto, a análise baseada apenas nas intensidade dos *voxels* pode produzir uma soluções pouco confiáveis, uma vez que ignora as considerações espaciais. Para contornar esse problema, o Atropos permite o uso de vários modelos de probabilidade *a priori*, incluindo MRF e informações baseadas em *templates* (imagens de referência com *voxels* rotulados para cada tecido). Neste trabalho, optou-se pelo uso do MRF, que modela as interações espaciais na vizinhança dos *voxels*. Com isso, o Atropos calcula a interação entre *voxels* considerando a distância Euclidiana entre eles, fazendo com que *voxels* mais próximos recebam um peso maior na análise do que aqueles mais distantes (AVANTS et al., 2011).

Antes da aplicação do Atropos, foi realizada a correção do *bias field* com o método “*N4 bias field correction*” da biblioteca ANTsPy para melhorar a precisão da segmentação. O algoritmo de K-means foi inicializado usando o agrupamento K-means com três classes de tecido ( $K = 3$ ), sendo utilizado como modelo *a priori* inicial para ajustar os parâmetros

iniciais do GMM. O K-means segmenta a imagem de entrada em 3 *clusters*, gerando uma imagem de probabilidade inicial para cada classe, em que cada *voxel* possui a probabilidade de pertencer a uma classe específica. Essa etapa inicial acelera a convergência do Atropos e evita mínimos locais que poderiam prejudicar a segmentação final.

A segmentação foi refinada por meio de otimização Expectation-Maximization (EM), com consistência espacial imposta através de um modelo de *Markov Random Field* (MRF). A intensidade da regularização espacial foi definida como 0,00001, e a vizinhança para cada voxel foi estabelecida como uma janela de  $1 \times 1 \times 1$  nas três dimensões espaciais. O algoritmo EM foi configurado para rodar no máximo 10 iterações, sem exigir uma melhoria mínima na verossimilhança entre elas. Essa primeira etapa produziu mapas probabilísticos para GM, WM e CSF. Esses mapas foram posteriormente usados como priors para uma segunda execução do Atropos. Nesta segunda etapa, a influência dos priors foi desativada, garantindo que a classificação se baseasse exclusivamente nas distribuições posteriores estimadas na fase inicial. Ambas as etapas de segmentação foram restritas à região intracraniana definida pela máscara cerebral do paciente.

O algoritmo *Expectation-Maximization* otimiza os parâmetros do GMM. Na etapa de “*Expectation*”, o GMM usa as intensidades dos *voxels* da imagem para calcular a probabilidade de cada *voxel* pertencer a cada uma das classes de tecido (componentes Gaussianas) a partir dos parâmetros atuais da GMM (médias, variâncias e pesos) (AVANTS et al., 2011). No passo de “*Maximization*”, o GMM atualiza seus parâmetros para maximizar a probabilidade das observações (intensidades dos *voxels*), dadas as probabilidades calculadas na etapa anterior. Para cada componente da GMM, a média é recalculada como a média ponderada das intensidades dos *voxels*, utilizando as probabilidades como pesos; a variância é recalculada com base nas intensidades dos *voxels* e nas médias atualizadas; o peso é atualizado de acordo com a proporção de *voxels* atribuídos a cada componente. Esse processo de otimização se repete até a convergência dos parâmetros.

A otimização no Atropos ocorre de forma síncrona e assíncrona. No esquema de otimização síncrona, o algoritmo percorre a imagem sem atualizar nenhum rótulo até que todos os *voxels* tenham sido visitados, e então atualiza todos os rótulos simultaneamente (AVANTS et al., 2011). Porém, como as atualizações síncronas nem sempre garantem convergência, o algoritmo também utiliza o esquema assíncrono, baseado no método *Iterated Conditional Modes* (BESAG, 1986). Esse método busca encontrar uma configuração de rótulos que maximiza a probabilidade posterior considerando a interação entre *voxels* vizinhos. Durante cada iteração, o algoritmo seleciona um rótulo (GM, WM ou CSF), percorre a imagem e avalia cada *voxel* na vizinhança com base em sua probabilidade de pertencer a cada rótulo, levando em conta tanto a intensidade do *voxel* quanto os rótulos dos *voxels* vizinhos. O rótulo de cada *voxel* é então atualizado para aquele que maximiza a probabilidade a posteriori, considerando a influência dos *voxels* adjacentes (AVANTS et al., 2011).

Como resultado, o Atropos gerou um mapa de segmentação rígida dos tecidos — atribuindo cada *voxel* a uma das três classes de tecido — e os mapas probabilísticos correspondentes. Essa abordagem em duas etapas melhorou a robustez da segmentação, reduziu o risco de convergência para mínimos locais e aumentou a plausibilidade anatômica dos mapas de tecido resultantes.

## 4.5 Parcelamento do cérebro em ROIs

Para extrair informações quantitativas de cada região cerebral, o atlas da NAC foi utilizado para segmentar o cérebro em ROIs. Uma vez que o alinhamento das imagens dos atlas com cada uma das imagens dos pacientes já foi executada (vide Subseção 4.3.5.3), cada imagem binária contendo uma região específica do cérebro foi sobreposta atuando como uma máscara para selecionar as ROIs a serem analisadas quantitativamente nesta pesquisa.

A seleção das ROIs a serem usadas nesta pesquisa foi feita com base nas regiões relevantes no contexto de DA explicitadas nos trabalhos correlatos (Seção 3.1) e na disponibilidade das malhas correspondentes a essas regiões no atlas da base NAC. Com isso, conforme Tabela 14, selecionou-se 22 regiões divididas em hemisférios esquerdo e direito, com quatro delas sendo divididas em 15 sub-regiões, totalizando 66 ROIs. Neste trabalho, as sub-regiões foram agrupadas por região formando uma única máscara para cada região, resultando em 44 ROIs no total (divididas em hemisférios esquerdo e direito).

Tabela 14 – Regiões segmentadas utilizadas nesta proposta de pesquisa.

Índice	Região	Índice	Região
1	Hipocampo	12	Giro fusiforme
2	Amígdala	13	Parietal
3	Giro temporal médio	14	Precuneus
4	Giro parahipocampal	15	Giro occipital lateral
5	Giro temporal inferior	16	Cuneus
6	Giro temporal superior	17	Giro pré-central
7	Caudado	18	Istmo do cíngulo
8	Ínsula	19	Pallidum
9	Putâmen	20	Giro frontal médio
10	Área de Accumbens	21	Giro cingulado anterior
11	Polo temporal	22	Tálamo

Nota 1: Regiões dos índices 1 ao 18 são divididas entre hemisfério esquerdo e direito.

Nota 2: Regiões dos índices 19 ao 21 inclui hemisfério esquerdo e direito de 2 sub-regiões diferentes.

Nota 3: Região de índice 22 inclui hemisfério esquerdo e direito de 9 sub-regiões diferentes.

Fonte: Autoria própria.

## 4.6 Extração de atributos

Para classificar as imagens de RM, foram utilizados atributos anatômicos como volume e espessura cortical, além de atributos de textura para capturar diferenças sutis entre os cérebros de CCLe e CCLp. Os atributos utilizados no trabalho são descritos a seguir.

### 4.6.1 Volume das ROIs

Para encontrar o volume das ROIs selecionadas, as máscaras das ROIs obtidas do atlas da NAC, já corrigidas com a imagem do paciente CCL, foram utilizadas para extrair o volume de cada região com base no número de *voxels* e no volume de cada *voxel*. Para evitar vieses na comparação entre pacientes, cada volume medido foi normalizado dividindo-o pelo volume total intracraniano do paciente.

Os volumes regionais do cérebro foram calculados com base nos rótulos do atlas NAC previamente mapeados para o espaço de imagem nativo de cada paciente (ver Subseção 4.3.5). Para cada ROI, o volume foi obtido contando o número de *voxels* atribuídos ao respectivo rótulo e multiplicando pelo volume do *voxel* isotrópico.

Para compensar as diferenças no tamanho da cabeça entre os pacientes, todos os volumes das ROIs foram normalizados pelo ICV individual. Além dos volumes totais por ROI, também foram calculados os volumes específicos por tecido — GM, WM e CSF — dentro de cada ROI, utilizando os mapas probabilísticos gerados pela segmentação Atropos (ver Seção 4.4). Esses volumes por tecido foram posteriormente normalizados pelo volume da ROI correspondente, resultando na composição relativa de cada tipo de tecido, permitindo assim uma caracterização morfométrica mais detalhada.

### 4.6.2 Espessura cortical

Para a extração da espessura cortical foi utilizado o algoritmo *Kelly Kapowski* do pacote ANTs, que implementa a técnica de Espessura Cortical Baseada em Registro Difeomórfico (*Diffeomorphic Registration-based Cortical Thickness*, DiReCT). Essa técnica calcula a distância entre as superfícies de substância branca e pia por meio da deformação difeomórfica dos mapas de probabilidade de substância cinzenta e branca obtidos na segmentação dos tecidos.

O algoritmo foi aplicado utilizando três entradas: a imagem segmentada de tecidos, juntamente com os mapas de probabilidade de substância cinzenta e branca gerados pela segmentação Atropos (Seção 4.4). A estimativa da espessura cortical foi realizada com dez iterações de otimização, um tamanho de passo de gradiente de 0,5 e um peso do modelo prior de 1 no processo de minimização de energia. Esses valores de parâmetros garantiram um equilíbrio entre eficiência computacional e precisão anatômica, promovendo uma convergência estável e preservando os detalhes estruturais.

Os mapas resultantes de espessura cortical foram parcelados utilizando o atlas NAC, permitindo a extração dos valores médios de espessura cortical dentro de cada região de interesse anatomicamente definida. Para compensar as diferenças no tamanho da cabeça entre os pacientes, as medidas de espessura cortical foram normalizadas pelo ICV individual.

### 4.6.3 Atributos de textura

Para complementar as informações volumétrica e de espessura, foram utilizados atributos de textura obtidos do GLCM (vide Subseção 2.3.1), uma vez que este método é amplamente utilizado nas pesquisas envolvendo uso de atributos de textura para detectar mudanças neurodegenerativas envolvidas na progressão da DA (LEE; LEE; KIM, 2020; SHU et al., 2021; SILVA et al., 2023; ZAIN; MAHMUD; GAN, 2023).

As GLCMs foram calculadas utilizando uma distância intervoxel fixa de 1 voxel, matrizes de coocorrência simétricas e uma largura de discretização de 64 níveis de cinza. Esses parâmetros foram escolhidos para garantir sensibilidade a heterogeneidades locais, mantendo a robustez ao ruído e a eficiência computacional. As medidas extraídas foram selecionadas com base nos achados de estudos recentes (LEE; LEE; KIM, 2020; MALLIKA et al., 2023), as principais fórmulas e descrições correspondentes estão apresentadas no Apêndice 7.1.

#### 4.6.3.1 Características de Textura do Cérebro Inteiro

Conforme mencionado na Subseção 2.3.2, a escolha do método de redução de ruído a ser aplicado nas imagens de RM antes da extração de atributos de textura pode afetar ligeiramente os valores desses atributos, podendo impactar na classificação das imagens de RM. Assim, os atributos de textura foram extraídos a partir de dois tratamentos diferentes de imagem: imagem original (sem redução de ruído) e imagem processada com NLM (SILVA et al., 2023; MALLIKA et al., 2023). Além disso, foram avaliados diferentes níveis de quantização na construção das matrizes GLCM (SHU et al., 2021). Diante disso, três variantes de pré-processamento foram consideradas:

1. **Tipo I:** imagens com remoção do crânio, reamostradas para resolução isotrópica de  $1 \text{ mm}^3$  e normalizadas em intensidade;
2. **Tipo II:** Tipo I com adição de redução de ruído e correção de campo de bias;
3. **Tipo III:** Tipo II com padronização de histograma adicional.

Para cada variante, as GLCMs foram calculadas usando tanto 32 quanto 64 bins de quantização de níveis de cinza. Conforme discutido na Subseção 5.2.7, as imagens do Tipo II com 64 bins produziram as características mais estáveis e informativas.

### 4.6.3.2 Características de Textura Baseadas em ROIs

As características de GLCM também foram extraídas de 66 ROIs em todas as configurações de pré-processamento e quantização. Em consonância com a análise de cérebro inteiro, as imagens do Tipo II com 64 bins demonstraram o melhor desempenho. O impacto da harmonização também foi avaliado comparativamente (ver Subseção 5.2.8).

## 4.7 Correção do efeito do *scanner*

Para corrigir o efeito do *scanner* e obter atributos radiômicos mais confiáveis e reproduzíveis, foi utilizado o método ComBat (vide Subseção 2.3.3) sobre todos os atributos extraídos. Utilizando o pacote `neuroComBat`<sup>4</sup> do Python, as imagens adquiridas por *scanners* 1,5T e 3T foram tratadas como dois lotes distintos e o grupo clínico (CCLp vs CCLe) foi incluído como covariável, a fim de preservar a variabilidade biologicamente relevante. Inicialmente, a distribuição de cada atributo em cada lote foi avaliada. Como a maioria das distribuições dos atributos extraídos não segue uma distribuição normal, o ComBat não paramétrico foi aplicado. Além disso, o método EB foi empregado para ajustar as estimativas de harmonização com base na distribuição de todos os atributos simultaneamente.

Tanto os conjuntos de características harmonizados quanto os não harmonizados foram mantidos para análises posteriores, permitindo uma avaliação explícita dos impactos da harmonização.

## 4.8 Modelo de classificação

Esta seção descreve o arcabouço de aprendizado de máquina utilizado para classificar os sujeitos como CCLe ou CCLp, com base nas características radiômicas extraídas. Serão detalhadas a estratégia de seleção de atributos adotada para lidar com a maldição da dimensionalidade, a justificativa para a escolha do modelo e os procedimentos utilizados para validar o desempenho do classificador.

### 4.8.1 Seleção de atributos

Três abordagens distintas foram consideradas para a seleção de atributos:

1. **Sem seleção de atributos:** Todas as características harmonizadas e filtradas foram mantidas sem redução de dimensionalidade, servindo como linha de base para comparação.

---

<sup>4</sup> <https://github.com/Jfortin1/ComBatHarmonization>

2. **Importância de atributos baseada em Floresta Aleatória:** Um classificador Floresta Aleatória foi treinado com validação cruzada aninhada (descrita abaixo), e os atributos foram ranqueados com base em seus escores de importância. Subconjuntos de atributos foram criados selecionando aqueles acima dos percentis 10, 25, 50, 75 e 90 da distribuição de importância. Essa estratégia utiliza o mecanismo interno de ranqueamento do modelo para orientar a redução de dimensionalidade.
3. **Seleção de atributos baseada em agrupamento hierárquico (*clustering*):** Para minimizar redundância e multicolinearidade, aplicou-se o agrupamento hierárquico aglomerativo com o método de ligação de Ward sobre a matriz de distâncias de correlação de Pearson absoluta (ou seja, 1 menos o coeficiente de correlação absoluto). Os *clusters* foram definidos cortando o dendrograma em um limiar de distância de 0,5, e o primeiro atributo de cada *cluster* foi retido.

### 4.8.2 Seleção do modelo de classificação

O número de atributos extraídos em comparação com o número de imagens disponíveis para os pacientes CCLe e CCLp é elevado, tornando a seleção de atributos um processo crítico para o desenvolvimento de um modelo robusto e eficaz.

A quantidade limitada de imagens de CCL disponíveis inviabiliza o uso de técnicas de aprendizado profundo, que requerem a estimação de muitos parâmetros, podendo resultar em um modelo sobreajustado. Por outro lado, técnicas convencionais, como SVM, regressão logística e árvores de decisão (SHU et al., 2021; XIAO et al., 2021; VELAZQUEZ; LEE, 2021) podem deixar de capturar relações importantes entre as características, ocasionando alto viés e baixa variância, levando a um subajuste.

Diante disso, foi utilizado um modelo baseado em Floresta Aleatória (vide Subseção 2.4.2) para atuar, concomitantemente, tanto na seleção de atributos quanto na classificação final das imagens de CCL.

A Floresta Aleatória possui diversas vantagens, como a capacidade de avaliar a importância dos atributos tanto no nível do modelo quanto para previsões individuais, além de oferecer interpretabilidade, operar como um classificador não paramétrico e lidar bem com dados faltantes e valores atípicos, além de capturar superfícies de decisão não lineares (ALI et al., 2012). Contudo, esse modelo pode apresentar alta complexidade computacional ao empregar um grande número de árvores e estar suscetível ao *overfitting* na presença de dados ruidosos.

### 4.8.3 Treinamento do modelo

Para evitar viés nos resultados do modelo, optou-se por utilizar apenas uma imagem de cada paciente por grupo, resultando em 272 imagens de pacientes CCLe e 256 imagens de pacientes CCLp da base de imagens ADNI, totalizando 528 imagens. Diante disso,

nota-se um desbalanceamento de classes, com CCLe representando 51,66% do conjunto e CCLp apenas 48,34%.

Dado o número limitado de dados e o desbalanceamento entre as classes, adotou-se a validação cruzada aninhada *Nested Cross-Validation* (NCV) com estratificação, utilizando o *StratifiedKFold* para formar os conjuntos de treino, validação e teste, de modo a preservar a proporção entre as classes em cada subdivisão. No *loop* externo, os dados foram divididos em 10 partes, mantendo a proporção de 51% para CCLe e 48% para CCLp. No *loop* interno, os dados que não forem selecionados para teste foram subdivididos em 5 partes, sendo uma para validação e as restantes para treinamento em cada iteração. Dentro do *loop* interno, foi realizado o ajuste de hiperparâmetros via busca exaustiva em grade (*grid search*).

Para o treinamento do modelo, foi utilizado o pacote *Scikit-learn*<sup>5</sup> do Python, que oferece uma implementação eficiente e de fácil uso para Florestas Aleatórias, além de várias ferramentas para avaliação de modelos.

Esse pacote utiliza como padrão o algoritmo *Classification and Regression Trees* (CART) para construir cada uma das árvores de decisão. O processo inicia-se com a construção máxima da árvore, utilizando um método de particionamento binário recursivo (de ALVARENGA JÚNIOR, 2018). Em cada divisão, os dados são separados em duas sub-regiões, e o algoritmo continua particionando essas sub-regiões de forma contínua até o ponto de parada determinado por parâmetros que controlam: a profundidade máxima de cada árvore, o número mínimo de amostras que um nó deve ter antes de ser dividido em dois nós-filhos e o número mínimo de amostras que deve estar presente em uma folha. Com o ajuste desses parâmetros, foi possível controlar a complexidade das árvores, evitando o sobreajuste aos dados de treinamento.

O classificador Floresta Aleatória teve duas classes potenciais para sua saída: CCLp e CCLe. Essas classes foram “votadas” por cada árvore individual, cujos resultados foram agregados para fornecer uma probabilidade geral de conversão para DA. Posteriormente, foi avaliada a importância de cada atributo incluído no modelo, conforme explicado na Subsubseção 2.4.2.1.

Os hiperparâmetros avaliados durante o ajuste incluíram:

- a) Número de árvores: 100 ou 200;
- b) Profundidade máxima das árvores: ilimitada, 10 ou 20 níveis;
- c) Número mínimo de amostras para dividir um nó interno: 2, 5 ou 10;
- d) Número mínimo de amostras em um nó folha: 1, 2 ou 5.

Esses valores foram selecionados com base em análises exploratórias prévias e heurísticas de regularização amplamente utilizadas em modelos baseados em árvores. Para cada configuração testada, o modelo que obteve o maior F1-score ponderado durante a vali-

<sup>5</sup> <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html>

dação no laço interno foi selecionado. Um *seed* aleatório fixo foi utilizado em todos os experimentos para garantir reprodutibilidade.

#### 4.8.4 Métricas de avaliação

Para fornecer uma avaliação robusta do desempenho do classificador diante de desequilíbrio entre classes, foram empregadas múltiplas métricas de avaliação. Entre elas: F1-score ponderado, que serviu como critério principal de seleção de modelo; F1-score; Área sob a curva ROC (AUC); Sensibilidade (Sens.); especificidade (Spec.); Acurácia balanceada (Bal. Acc.). Essas métricas foram calculadas a partir dos verdadeiros positivos (VP), dos falsos positivos (FP), dos verdadeiros negativos (VN) e dos falsos negativos (FN).

Nesta pesquisa, o VP foi definido como o número de pacientes que realmente progrediram para DA e que o modelo previu corretamente como CCLp. O FP foi definido como o número de pacientes que não progrediram para DA, mas que o modelo classificou erroneamente como CCLp. O VN foi definido como o número de pacientes que não progrediram para DA e que o modelo previu corretamente como CCLe. Por fim, o FN foi definido como o número de pacientes que realmente progrediram para DA, mas que o modelo classificou erroneamente como CCLe.

A partir dos quatro valores obtidos no teste de classificação, a sensibilidade calcula a porcentagem de acertos na classe positiva, enquanto a especificidade mede a porcentagem de acertos na classe negativa. Essas métricas podem ser expressas pelas Equações 16 e 17:

$$\text{Sens} = \frac{\text{VP}}{\text{VP} + \text{FN}}, \quad (16)$$

$$\text{Espec} = \frac{\text{VN}}{\text{VN} + \text{FP}}. \quad (17)$$

A acurácia, que mede a quantidade total de acertos do classificador, é definida pela Equação 18 e a acurácia ponderada pela Equação 19:

$$\text{Acc} = \frac{\text{VP} + \text{VN}}{\text{VP} + \text{FP} + \text{VN} + \text{FN}}, \quad (18)$$

$$\text{Acc balanceada} = \frac{\text{Sens} + \text{Espec}}{2}. \quad (19)$$

Já o F1-score calcula a média harmônica entre precisão (Equação 20) e sensibilidade (Equação 16) é dada pela Equação 21:

$$\text{Precisão} = \frac{\text{VP}}{\text{VP} + \text{FP}}, \quad (20)$$

$$\text{F1-score} = \frac{2 \times \text{Precisão} \times \text{Sens}}{\text{Precisão} + \text{Sens}}. \quad (21)$$

Por fim, o F1-score ponderado é a média ponderada dos F1-scores de cada classe, levando em conta o número de amostras de cada classe, conforme Equação 22:

$$\text{F1-score ponderado} = \sum_{i=1}^C \frac{n_i}{N} F1_i, \quad (22)$$

sendo:

- a)  $c$  é número total de classes (no caso são duas: CCLp e CCLe),
- b)  $n_i$  é número de amostras da classe  $i$ ,
- c)  $N$  é número total de amostras.

## Avaliação da importância dos atributos

Para analisar a relevância dos atributos na previsão do modelo, utilizou-se a abordagem SHAP<sup>6</sup> (SHapley Additive exPlanations) (LUNDBERG; LEE, 2017), com a ferramenta TreeExplainer. Esse algoritmo calcula a importância dos atributos para a previsão do modelo com base em valores exatos de Shapley em tempo polinomial, tornando o processo mais eficiente e viável para modelos baseados em árvores, ao mesmo tempo que garante interpretabilidade (LUNDBERG; LEE, 2017).

Para isso, ele realiza o cálculo de explicações locais, em que cada valor de Shapley é calculado com base no ganho médio da previsão quando uma característica específica é incluída no modelo em comparação com quando ela é excluída (LUNDBERG; LEE, 2017). No contexto das árvores de decisão, isso é feito com base na estrutura da árvore (como as divisões nas folhas) e na forma como as características influenciam essas divisões, calculando as contribuições de forma mais eficiente (LUNDBERG; LEE, 2017).

Após calcular os valores de Shapley para todos os atributos, podemos interpretar como cada atributo contribui para a previsão do modelo. A soma de todos os valores de Shapley para uma instância será igual à diferença entre a previsão do modelo para aquela instância e a média das previsões do modelo (o valor base) (LUNDBERG; LEE, 2017). Isso fornece uma explicação clara e quantitativa da contribuição de cada característica para a previsão.

Os valores de Shapley, originados da teoria dos jogos, são amplamente utilizados para atribuição de importância dos atributos e satisfazem três propriedades fundamentais (LUNDBERG; LEE, 2017). A precisão local garante que a soma das contribuições atribuídas a cada atributo do modelo resulte exatamente na previsão feita para uma determinada entrada. A consistência estabelece que, se a contribuição de um atributo para a previsão aumentar em um modelo modificado, seu valor de Shapley não deve diminuir. Por fim, a ausência de efeito assegura que características sem influência na previsão recebam atribuição nula.

<sup>6</sup> <https://shap.readthedocs.io/en/latest/index.html>

Além disso, o TreeExplainer amplia as explicações locais ao capturar efeitos de interação entre variáveis por meio dos valores de interação SHAP, que utilizam o índice de interação de Shapley da teoria dos jogos (LUNDBERG; LEE, 2017). Essa abordagem distribui o crédito entre variáveis individuais e pares de variáveis, representando as interações em uma matriz. Isso permite uma análise mais detalhada das previsões do modelo, revelando padrões significativos.

## 4.9 Considerações finais

Este capítulo apresentou as bases de dados utilizadas para o desenvolvimento da pesquisa proposta, além de detalhar as técnicas de pré-processamento, segmentação e extração de atributos radiômicos, bem como os métodos para mitigar os efeitos do *scanner*. Também foram discutidos aspectos da metodologia relacionados ao desenvolvimento do modelo de classificação e às métricas de avaliação dos modelos. No Capítulo 5 serão apresentados os resultados e discussões provenientes desta pesquisa.

---

# Capítulo 5

## Resultados e Discussões

---

*Este capítulo apresenta e discute os resultados obtidos nesta pesquisa, como o pré-processamento das imagens, atributos extraídos e os experimentos de classificação das imagens de RM em CCLe e CCLp.*

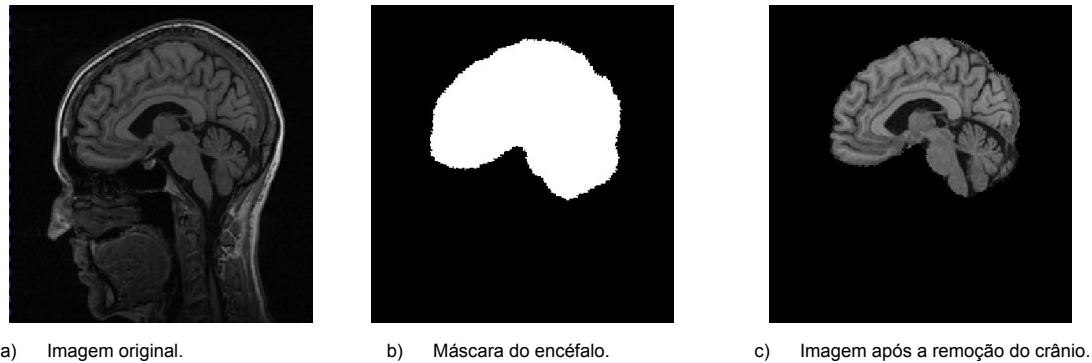
### 5.1 Pré-processamento

As imagens de RM empregadas nesta pesquisa foram submetidas a uma sequência de etapas de pré-processamento que incluiu a remoção do crânio, redução de ruído, a correção do *bias field*, padronização de histogramas, o alinhamento espacial (corregistro) com o atlas e a segmentação das 44 ROIs (definidas na Subseção 4.5). Devido ao grande volume de imagens, não foi possível inspecionar visualmente todas elas após o processamento. A seguir são apresentados alguns exemplos de imagens resultantes de cada etapa.

#### 5.1.1 Remoção do crânio

A Figura 26 (a) apresenta um exemplo de uma imagem da ADNI original (com o crânio e outros tecidos não-cerebrários), enquanto a Figura 26 (b) exhibe a máscara do encéfalo, a qual corresponde a área da região mantida na imagem, conforme mostrada na Figura 26 (c). Comparando-se as imagens, antes e após a remoção do crânio, nota-se que o algoritmo ROBEX (IGLESIAS et al., 2011) foi capaz de preservar a massa cerebral, além do cerebelo e o tronco encefálico, de forma eficaz e sem causar perdas significativas de informações estruturais na imagem. Portanto, pode-se dizer que o algoritmo teve resultados satisfatórios para o propósito deste trabalho.

Figura 26 – Resultado da remoção do crânio em uma das imagens selecionadas da ADNI.



Fonte: Autoria própria.

### 5.1.2 Redução de ruído

Na primeira linha da Figura 27, tem-se um exemplo de uma imagem da ADNI com corte sagital obtida da etapa de remoção do crânio (com ruído) (primeira imagem), o ruído removido (segunda imagem) e o resultado obtido da redução de ruído realizada pelo algoritmo NLM (terceira imagem). Já na segunda linha da Figura 27, tem-se uma região ampliada da imagem da ADNI com corte axial (primeira imagem), o ruído removido (segunda imagem) e o resultado obtido da redução de ruído realizada pelo algoritmo NLM (terceira imagem). Comparando-se as imagens, antes e após a redução do ruído, nota-se que o algoritmo NLM da biblioteca ANTsPy foi capaz de reduzir o ruído sem causar perdas significativas de informações estruturais na imagem. Portanto, pode-se dizer que o algoritmo teve resultados satisfatórios para o propósito deste trabalho.

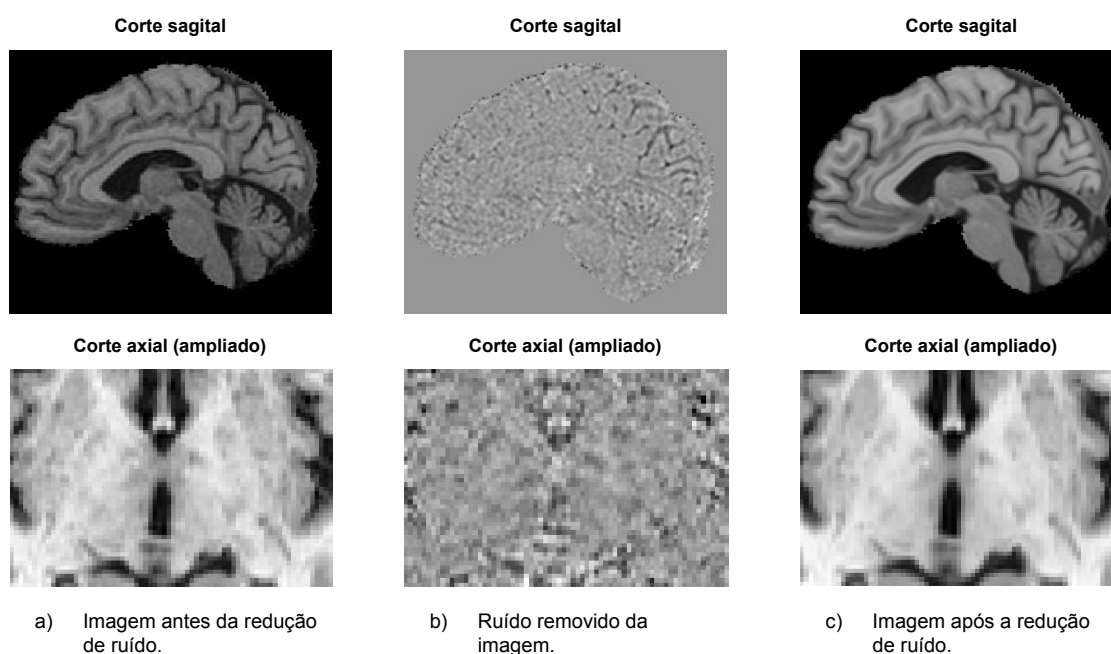
### 5.1.3 Correção de *bias field*

Nesta etapa de pré-processamento, foi possível observar que as imagens adquiridas da ADNI não são suficientemente afetadas pelo *bias field* para ser possível observar diferenças visuais significativas entre as imagens de entrada e saída da correção de *bias field*, respectivamente mostradas nas Figuras 28 (a) e (c). Entretanto, é possível verificar a eficácia do algoritmo pelo respectivo *bias field* estimado que foi corrigido utilizando o algoritmo “*N4 bias field correction*” da biblioteca ANTsPy mostrado na Figura 28 (b).

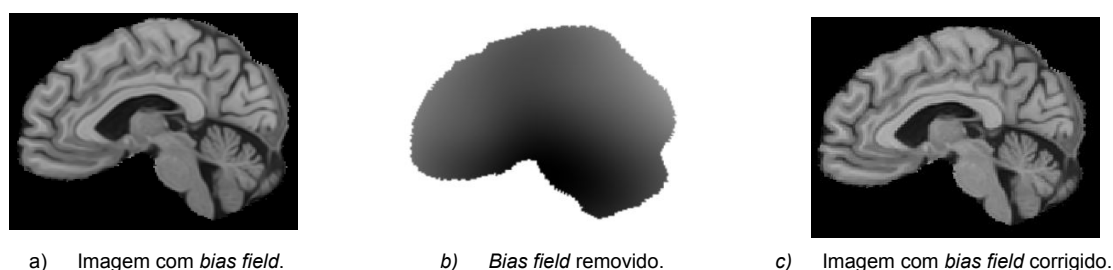
### 5.1.4 Padronização de histograma

Para a etapa de padronização de histograma, a imagem de referência da NAC e as imagens da ADNI foram normalizadas previamente, obtendo-se os histogramas mostrados na Figura 29(a) e (b), respectivamente, sendo (b) o histograma de uma das imagens da ADNI. Com isso, ao aplicar o método “*Histogram Matching Image Filter*” utilizando a

Figura 27 – Resultado da redução de ruído em uma das imagens selecionadas da ADNI.



Fonte: Autoria própria.

Figura 28 – Resultado da correção de *bias field* em uma das imagens selecionadas da ADNI.

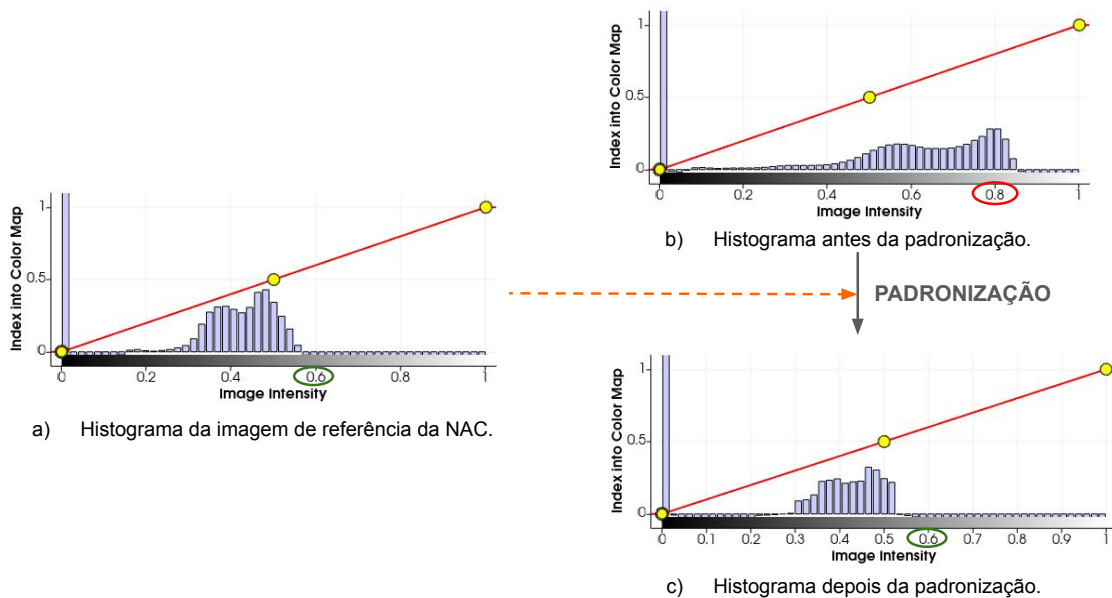
Fonte: Autoria própria.

biblioteca SimpleITK (NYÚL; UDUPA; ZHANG, 2000), foi possível observar um ajuste efetivo dos histogramas das imagens selecionadas da ADNI. Após a correspondência de histogramas, as distribuições de intensidades passaram a variar no intervalo de 0 a 0,6, conforme mostrado na Figura 29(c).

### 5.1.5 Corregistro

A técnica de “tabuleiro de xadrez” (*checkerboard*), que gera uma imagem de saída a partir da intercalação consecutiva de blocos das duas imagens, foi usada para permitir uma análise qualitativa dos corregistros produzidos utilizando transformações rígida, afim

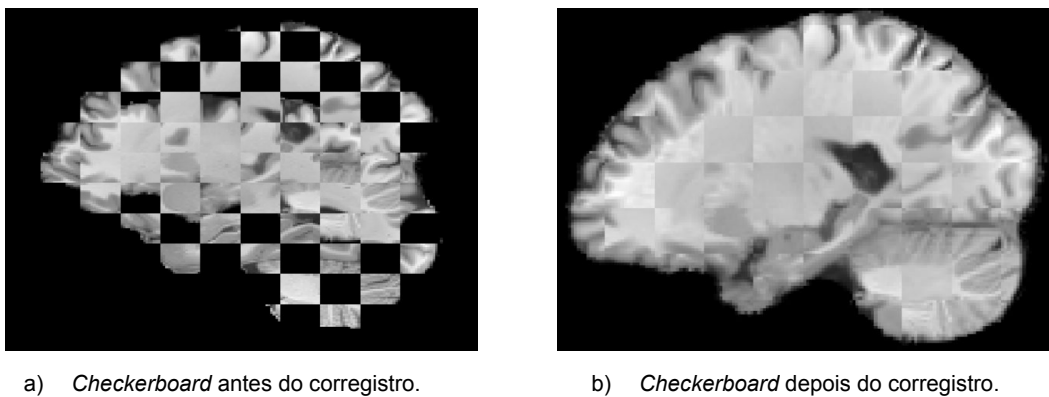
Figura 29 – Resultado da padronização de histograma de uma das imagens selecionadas da ADNI com a imagem de referência da NAC.



Fonte: Autoria própria.

e *Symmetric Normalization*. Na Figura 30, são apresentados os *checkerboards* compostos por uma imagem da ADNI e a imagem de referência da NAC antes (a) e depois do corregristo (b). Analisando os resultados, nota-se que as transformações aplicadas foram efetivas para alinhar a imagem de referência da NAC às imagens da ADNI.

Figura 30 – Resultado do corregristo da imagem de referência da NAC com uma imagem da ADNI.



Fonte: Autoria própria.

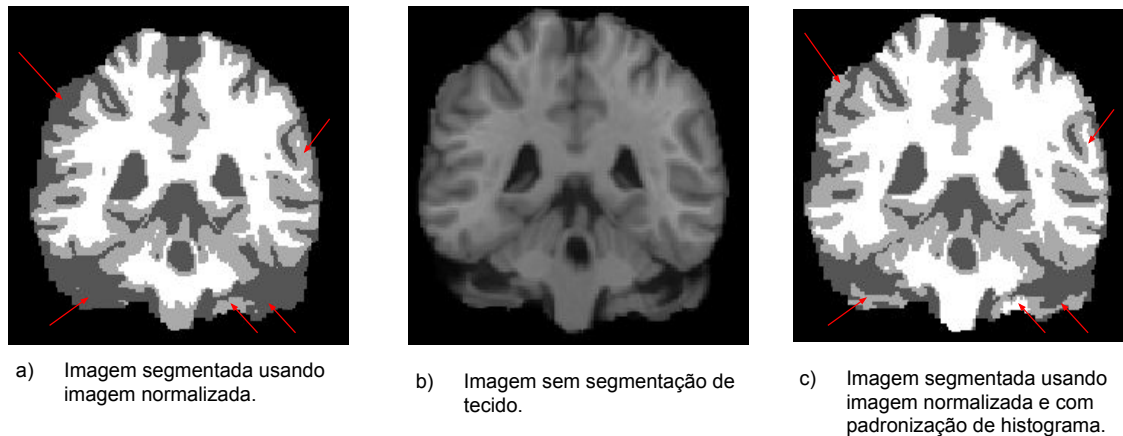
### 5.1.6 Segmentação dos tecidos cerebrais

A Figura 31 exibe uma imagem selecionada da ADNI submetida ao processo de segmentação de tecidos cerebrais, realizada pelo algoritmo Atropos do pacote ANTs. Na Figura 31(b), tem-se a imagem sem a segmentação de tecidos, enquanto em (a), tem-se a imagem resultante do processo de segmentação utilizando como entrada a imagem da ADNI normalizada e, em (c), a imagem segmentada a partir da imagem da ADNI normalizada e com histograma padronizado.

Como podemos observar visualmente, a segmentação realizada pelo algoritmo Atropos se mostrou mais eficaz ao utilizar como entrada a imagem da ADNI normalizada e com padronização de histograma (Figura 31(c)). Isso pode ser evidenciado principalmente nas regiões corticais e do cerebelo (indicada pelas setas vermelhas), uma vez que em (a) observa-se uma perda significativa das informações de tecidos de WM e GM quando comparado à imagem original, enquanto em (c) esses tecidos foram melhor preservados.

Diante disso, para a extração de atributos relacionados aos tecidos, utilizou-se as imagens segmentadas a partir das imagens da ADNI normalizadas e com padronização de histograma.

Figura 31 – Resultado da segmentação de tecidos cerebrais em uma imagem da ADNI normalizada com e sem padronização de histogramas.



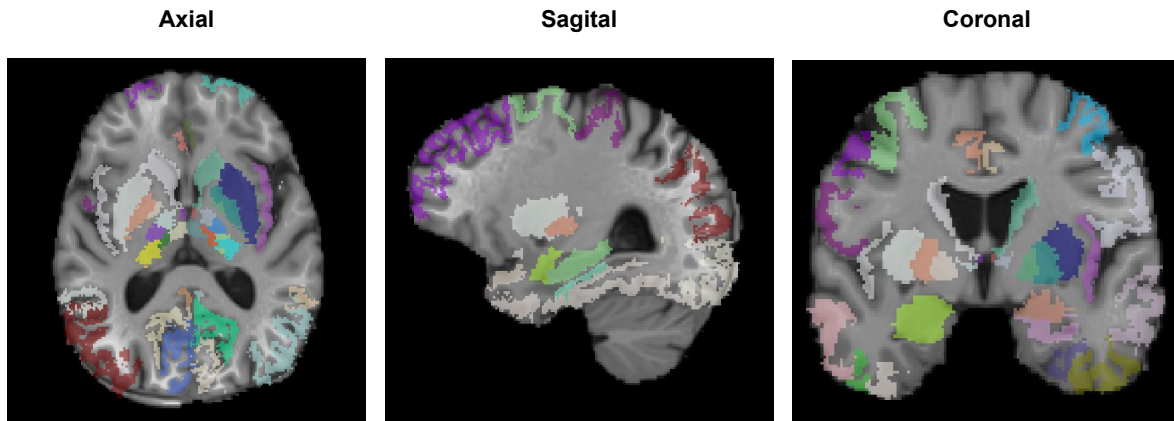
Fonte: Autoria própria.

### 5.1.7 Segmentação das regiões de interesse

A Figura 32 ilustra as máscaras das ROIs sobrepostas a uma imagem selecionada da ADNI. As máscaras foram utilizadas para definir as regiões de interesse nas imagens dos pacientes para a extração de medidas de espessura cortical, volume e textura em cada região potencialmente relacionada com a progressão de CCL para DA. As regiões segmentadas (regiões das máscaras) correspondem a 22 regiões divididas em hemisférios esquerdo e direito, totalizando 44 ROIs (definidas na Subseção 4.5).

Como podemos observar, as regiões de giros cerebrais (presentes no córtex cerebral), como o giro frontal médio e giro pré-central não foram perfeitamente alinhadas com a imagem da ADNI. Porém, optou-se por manter essas regiões para posteriores análises estatísticas das métricas extraídas.

Figura 32 – Resultado da segmentação das regiões de interesse em uma imagem da ADNI (vistas axial, sagital e coronal).



Fonte: Autoria própria.

## 5.2 Extração de atributos

Para desenvolver um modelo de classificação, baseado em uma assinatura radiômica, para prever a progressão de CCL para DA, foram extraídos atributos de volume, espessura cortical e textura de diferentes regiões cerebrais. Conforme apresentado na Tabela 15, foram extraídos 29 atributos de cada uma das 44 ROIs e 26 atributos da imagem toda, totalizando em 1302 atributos. Foram realizados experimentos utilizando tanto os dados originais quanto os dados harmonizados pelo ComBat, com o objetivo de comparar os efeitos da harmonização. Os resultados das análises estatísticas serão apresentados a seguir.

### 5.2.1 Volume total intracraniano

O volume total intracraniano corresponde ao volume da máscara obtida da etapa de remoção do crânio e reamostrada para uma resolução isotrópica de  $1\text{mm}^3$ . A Figura 33 exibe as distribuições desse atributo para as classes CCLe e CCLp, na qual é possível observar a sobreposição das distribuições entre as duas classes.

Inicialmente, foi realizado o teste de normalidade de Shapiro-Wilk, que indicou que os dados não seguem uma distribuição normal ( $p\text{-valor} < 0,05$ ). Em seguida, foi realizado o teste de Welch (estatística =  $-1,0252$ ;  $p\text{-valor} = 0,3058$ ), cujos resultados corroboram a

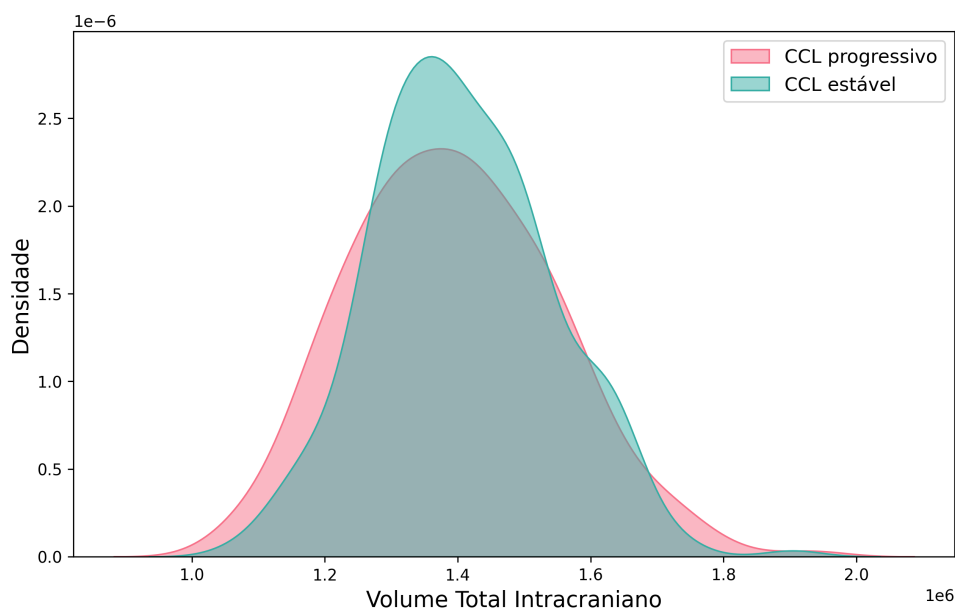
Tabela 15 – Conjuntos de atributos extraídos e suas quantidades.

Conjuntos de atributos extraídos	Qtde. total de métricas extraídas	
	De cada ROI	De toda a imagem
Volume total intracraniano	-	1
Volume de cada ROI	1	44
Volume de cada tecido (GM, WM e CSF) em cada ROI	3	132
Espessura cortical do cérebro todo	-	1
Espessura cortical em cada ROI	1	44
Atributos de textura na imagem toda	-	24
Atributos de textura em cada ROI	24	1056
<b>Total</b>	<b>29</b>	<b>1302</b>

Fonte: Autoria própria.

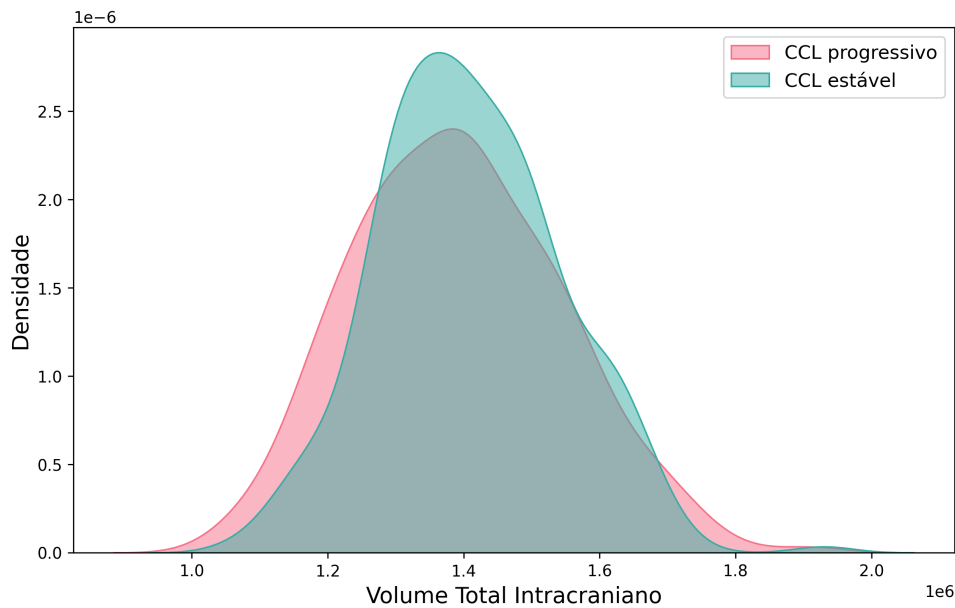
hipótese de que não há diferença significativa entre os volumes totais intracranianos das classes comparadas. Esse resultado também foi observado para os dados harmonizados com Combat, exibidos na Figura 34. Esses dados apresentaram um aumento na estatística de Welch (estatística = -1,3818; p-valor = 0,1676), indicando uma possível redução da variabilidade dos dados, mas ainda assim sem alcançar significância estatística para diferenciar as classes.

Figura 33 – Distribuição do volume total intracraniano entre as classes CCLe e CCLp (sem harmonização).



Fonte: Autoria própria.

Figura 34 – Distribuição do volume total intracraniano entre as classes CCLe e CCLp (com harmonização).



Fonte: Autoria própria.

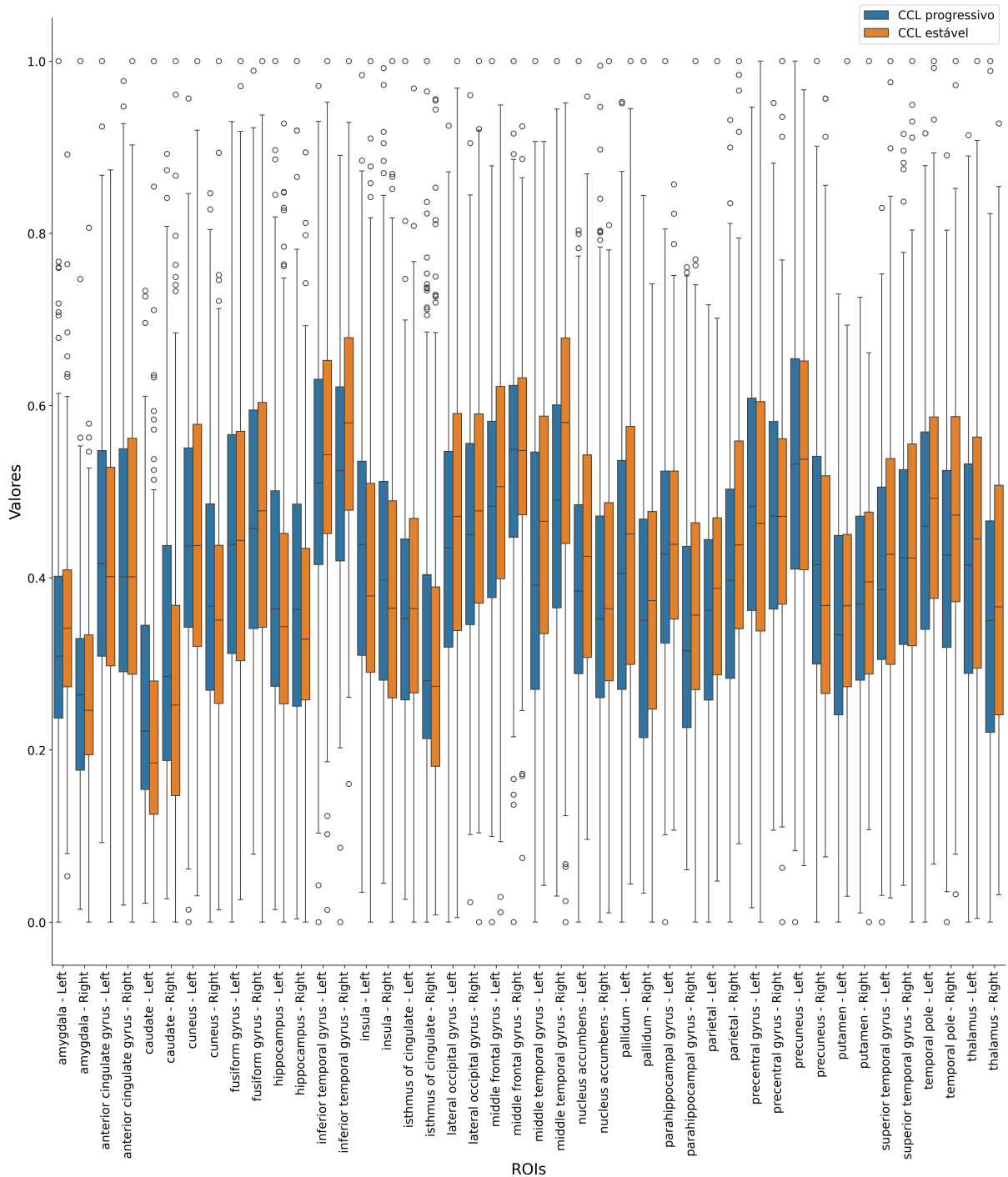
## 5.2.2 Volume das ROIs

### 5.2.2.1 Sem harmonização ComBat

Os volumes extraídos das 44 ROIs foram normalizados com o volume total intracraniano de cada paciente. A Figura 35 apresenta os *boxplots* das medidas de volume extraídas de cada ROI para as classes CCLe e CCLp, com normalização por coluna para facilitar a visualização. Visualmente, observa-se que os *boxplots* estão amplamente sobrepostos para a maioria dos atributos, com diferenças sutis na dispersão e com algumas diferenças expressivas nas medianas, como as da ínsula esquerda e do giro temporal médio (hemisférios esquerdo e direito).

Para uma análise quantitativa, foi realizado o teste de normalidade de Shapiro-Wilk, que indicou que 12 atributos seguem uma distribuição normal ( $p\text{-valor} > 0,05$ ). Em seguida, para esses atributos, foi aplicado o teste de Levene, que confirmou a homogeneidade das variâncias (homoscedasticidade). Para verificar a significância das diferenças entre as classes CCLe e CCLp, foi utilizado o teste t de Student nos atributos com distribuição normal e homoscedasticidade, revelando que 5 deles apresentaram diferenças estatisticamente significativas. Para os demais atributos, foi aplicado o teste de Welch, que indicou diferenças significativas em 13 atributos. Dessa forma, obteve-se um total de 18 atributos com diferenças estatisticamente significativas entre as classes CCLe e CCLp.

Figura 35 – Distribuição do volume de cada ROI nas classes CCLe e CCLp (sem harmonização).



Fonte: Autoria própria.

### 5.2.2.2 Com harmonização ComBat

Para uma análise quantitativa dos atributos após a harmonização ComBat, foi realizado o teste de normalidade de Shapiro-Wilk, que indicou que 15 atributos seguem uma distribuição normal ( $p\text{-valor} > 0,05$ ). Em seguida, para esses atributos, foi aplicado o teste

de Levene, que confirmou a homogeneidade das variâncias (homoscedasticidade). Para verificar a significância das diferenças entre as classes CCLe e CCLp, foi utilizado o teste t de Student nos atributos com distribuição normal e homoscedasticidade, revelando que 4 deles apresentaram diferenças estatisticamente significativas. Para os demais atributos, foi aplicado o teste de Welch, que indicou diferenças significativas em 9 atributos. Dessa forma, obteve-se um total de 13 atributos com diferenças estatisticamente significativas entre as classes CCLe e CCLp.

Conforme apresentado na Tabela 16, após a harmonização, a quantidade de atributos com significância estatística caiu de 18 para 13, incluindo os volumes do hipocampo (dos hemisférios esquerdo e direito) e do pallidum (hemisfério direito).

Tabela 16 – ROIs com volumes que apresentaram diferenças estatisticamente significativas entre CCLe e CCLp (com harmonização).

Sem harmonização	Com harmonização
Giro temporal inferior (direito)	Giro temporal inferior (direito)
Giro temporal médio (direito)	Giro temporal médio (direito)
Caudado (esquerdo)	Caudado (esquerdo)
Caudado (direito)	Caudado (direito)
Giro temporal inferior (esquerdo)	Giro temporal inferior (esquerdo)
Parietal (direito)	Parietal (direito)
Giro temporal superior (esquerdo)	Giro temporal superior (esquerdo)
Giro temporal médio (esquerdo)	Giro temporal médio (esquerdo)
Polo temporal (direito)	Polo temporal (direito)
Giro parahipocampal (direito)	Giro parahipocampal (direito)
Giro occipital lateral (esquerdo)	Hipocampo (esquerdo)
Giro frontal médio (esquerdo)	Hipocampo (direito)
Amígdala (esquerdo)	Pallidum (direito)
Ínsula (esquerdo)	
Área de Accumbens (esquerdo)	
Polo temporal (esquerdo)	
Parietal (esquerdo)	
Putâmen (esquerdo)	

Fonte: Autoria própria.

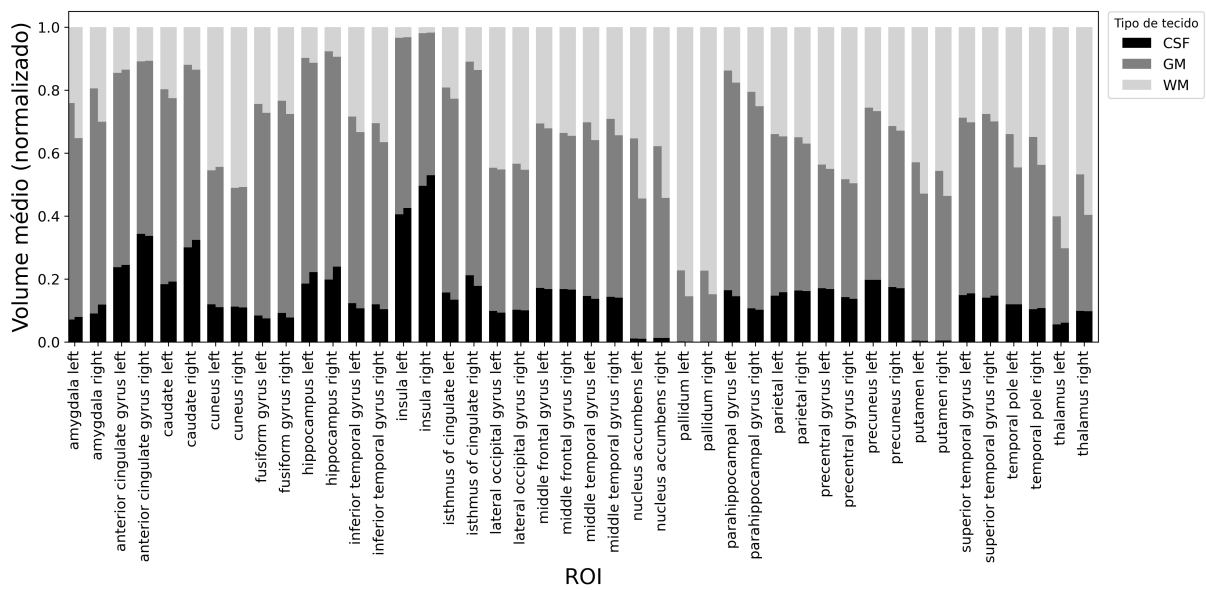
### 5.2.3 Volume de cada tecido

#### 5.2.3.1 Sem harmonização ComBat

Os volumes de cada tecido cerebral (CSF, GM e WM) extraídos de cada uma das 44 ROIs foram normalizados com o volume total de cada ROI. A Figura 36 apresenta, para cada classe (CCLe e CCLp), os histogramas empilhados com a média dos volumes de CSF, GM e WM extraídas de cada ROI. Nas barras de cada ROI, a primeira barra representa os volumes de CCLe, enquanto a segunda barra corresponde aos volumes de

CCLp. Visualmente, observa-se que o volume médio de GM é mais elevada na classe CCLe para a maioria das ROIs, com destaque para amígdala e a Área de Accumbens (nos hemisférios direito e esquerdo). Em relação a WM, a média do volume desse tecido é mais elevada na classe CCLp para a maioria das ROIs. Além disso, nota-se que, quando há uma redução no volume de GM na classe CCLp em comparação ao CCLe, essa perda parece ser compensada por um aumento no volume de WM.

Figura 36 – Volumes médios de CSF, GM e WM em cada ROI para as classes CCLe (primeira barra) e CCLp (segunda barra) (sem harmonização).



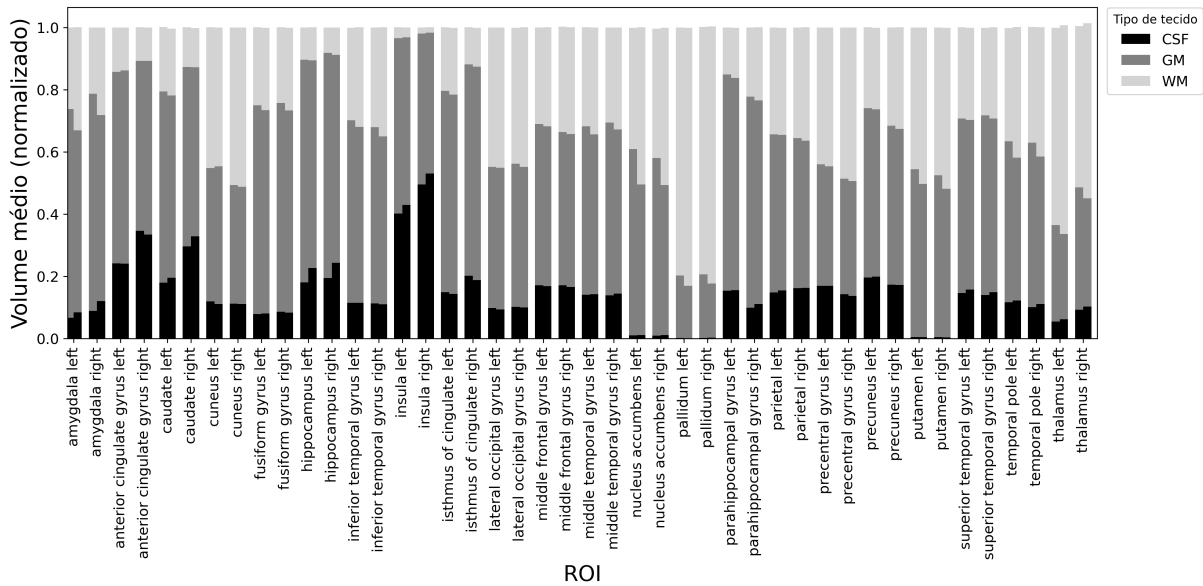
Fonte: Autoria própria.

Para uma análise quantitativa, foi realizado o teste de normalidade de Shapiro-Wilk, que indicou que 22 atributos seguem uma distribuição normal ( $p$ -valor  $> 0,05$ ). Em seguida, foi aplicado o teste de Levene para verificar a homogeneidade das variâncias, o qual confirmou a homoscedasticidade para 18 atributos e rejeitou para 4. Para verificar a significância das diferenças entre as classes CCLe e CCLp, utilizou-se o teste  $t$  de Student nos atributos com distribuição normal e homoscedasticidade, revelando que 12 deles apresentaram diferenças estatisticamente significativas. Para os atributos restantes, aplicou-se o teste de Welch, que identificou diferenças significativas em 71 atributos. Assim, obteve-se um total de 83 atributos com diferenças estatisticamente significativas entre as classes CCLe e CCLp, sendo 35 relativos ao volume de GM, 34 a WM e 14 a CSF.

### 5.2.3.2 Com harmonização ComBat

A Figura 37 apresenta, para cada classe (CCLe e CCLp), os histogramas empilhados com a média dos volumes de CSF, GM e WM extraídas de cada ROI após a harmonização ComBat. É possível observar sutis alterações, reduzindo a diferença entre as duas classes dentro de cada ROI.

Figura 37 – Volumes médios de CSF, GM e WM em cada ROI para as classes CCLe (primeira barra) e CCLp (segunda barra) (com harmonização).



Fonte: Autoria própria.

Para uma análise quantitativa dos atributos após a harmonização ComBat, foi realizado o teste de normalidade de Shapiro-Wilk, que indicou que 31 atributos seguem uma distribuição normal ( $p\text{-valor} > 0,05$ ). Em seguida, foi aplicado o teste de Levene para verificar a homogeneidade das variâncias, o qual confirmou a homoscedasticidade para 27 atributos e rejeitou para 4. Para verificar a significância das diferenças entre as classes CCLe e CCLp, utilizou-se o teste t de Student nos atributos com distribuição normal e homoscedasticidade, revelando que 11 deles apresentaram diferenças estatisticamente significativas. Para os atributos restantes, aplicou-se o teste de Welch, que identificou diferenças significativas em 50 atributos. Assim, obteve-se um total de 61 atributos com diferenças estatisticamente significativas entre as classes CCLe e CCLp, sendo 31 relativos ao volume de GM, 18 a WM e 12 a CSF.

Com a harmonização, o número de atributos estatisticamente significativos para a diferenciação entre as classes reduziu de 83 para 61, mantendo a predominância dos tecidos GM, seguidos por WM e CSF. Além disso, 54 atributos permaneceram significativos, enquanto 29 perderam a significância e 7 novos atributos passaram a apresentar significância.

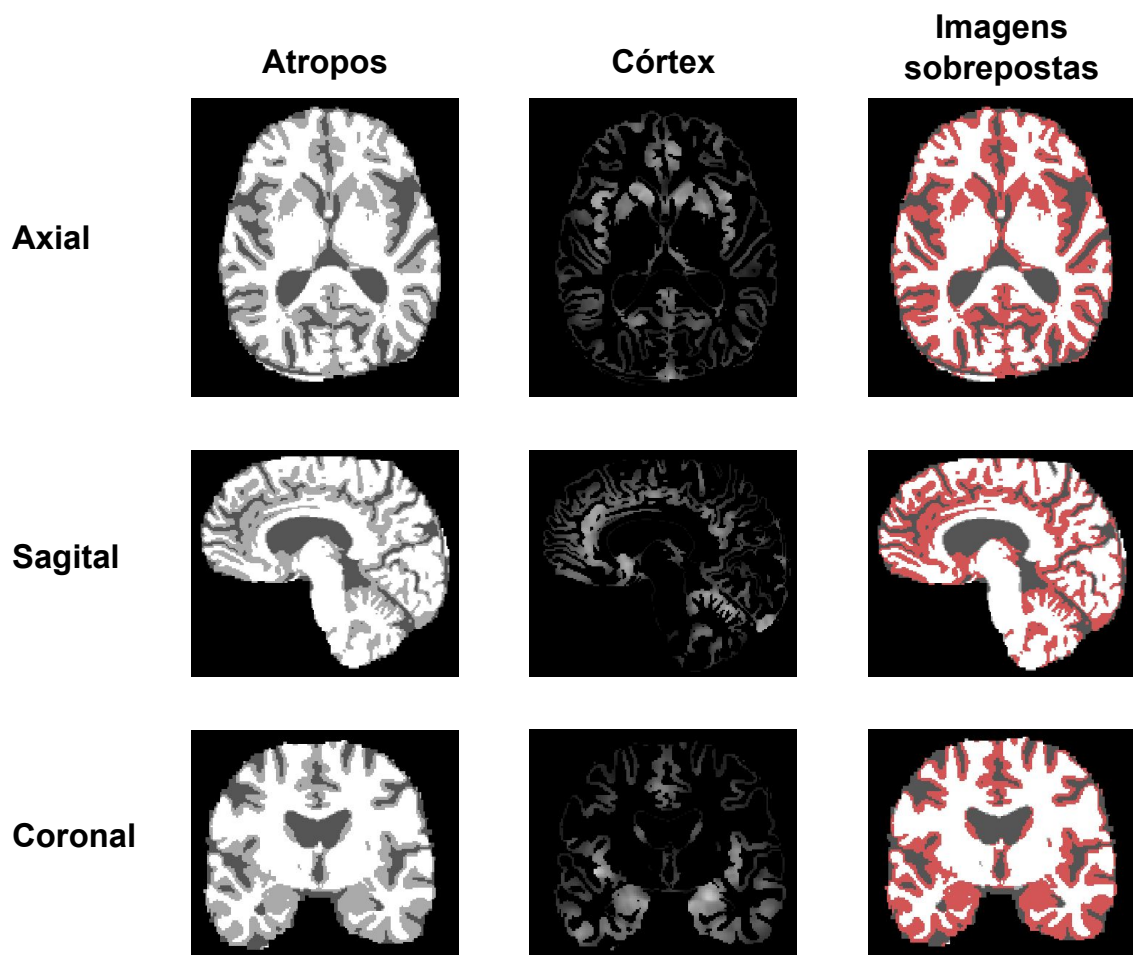
#### 5.2.4 Imagem da espessura cortical

Para extrair as medidas de espessura cortical foi necessário, primeiramente, obter a imagem da espessura cortical a partir do algoritmo Kelly Kapowski do pacote ANTs. A Figura 38 exibe o resultado da geração da imagem de espessura cortical a partir de uma imagem selecionada da ADNI. Nesta figura, são apresentadas as visualizações axial

(primeira linha), sagital (segunda linha) e coronal (terceira linha) para três tipos de imagens: as imagens resultantes do Atropos (primeira coluna), que foram utilizadas como entrada juntamente com as imagens de probabilidade; a imagem da espessura cortical gerada (segunda coluna); e a sobreposição de ambas as imagens (terceira coluna), com a espessura cortical destacada em vermelho.

Analisando os resultados, nota-se que o algoritmo Kelly Kapowski foi efetivo para obter as imagens de espessura cortical a partir das imagens da ADNI com tecidos segmentados, apresentando resultados qualitativos satisfatórios para a posterior extração de medidas de espessura cortical.

Figura 38 – Resultado da geração da imagem de espessura cortical a partir de uma imagem da ADNI (vistas axial, sagital e coronal).



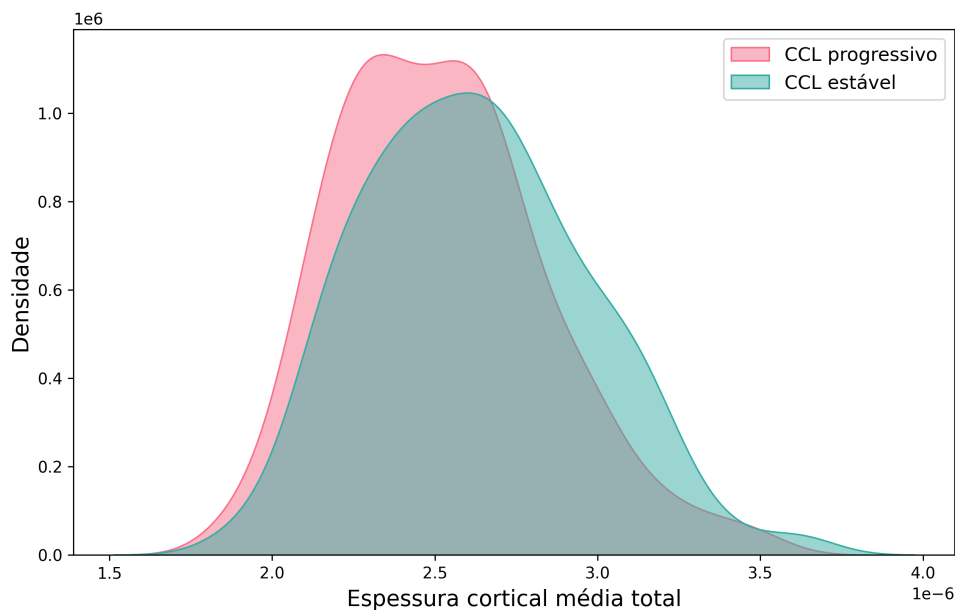
Fonte: Autoria própria.

## 5.2.5 Espessura cortical total

### 5.2.5.1 Sem harmonização ComBat

A espessura cortical média total foi extraída considerando todo o córtex do encéfalo e normalizada com o volume total intracraniano de cada paciente. A Figura 39 exhibe as distribuições desse atributo para as classes CCLe e CCLp, na qual é possível observar que a curva da classe CCLp está levemente deslocada para a esquerda em relação a curva da classe CCLe, apesar de apresentarem uma grande área sobreposta.

Figura 39 – Distribuição da espessura cortical média total entre as classes CCLe e CCLp (sem harmonização).



Fonte: Autoria própria.

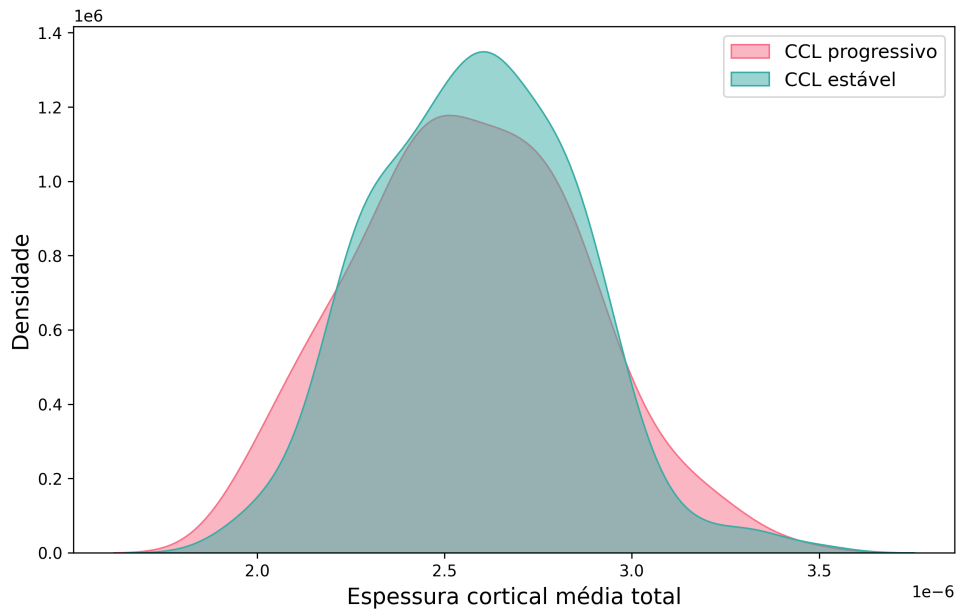
Inicialmente, para uma análise quantitativa, foi realizado o teste de normalidade de Shapiro-Wilk, que indicou que os dados não seguem uma distribuição normal ( $p$ -valor  $< 0,05$ ). Em seguida, foi realizado o teste de Welch (estatística =  $-3,6126$ ;  $p$ -valor =  $0,0003$ ), indicando que há diferenças estatisticamente significativas entre as espessuras corticais médias totais entre as classes CCLp e CCLe.

### 5.2.5.2 Com harmonização ComBat

Para uma análise quantitativa dos atributos após a harmonização ComBat, inicialmente, foi realizado o teste de normalidade de Shapiro-Wilk, que indicou que os dados seguem uma distribuição normal ( $p$ -valor  $> 0,05$ ). Em seguida, foi aplicado o teste de Levene, que rejeitou a homogeneidade do atributo. Em seguida, foi realizado o teste de Welch (estatística =  $-0,6273$ ;  $p$ -valor =  $0,5307$ ), indicando que não há diferenças estatisticamente significativas entre as espessuras corticais médias totais entre as classes CCLp

e CCLe. Isso pode ser observado visualmente conforme apresentado na Figura 40, em que se observa uma maior sobreposição das distribuições das espessuras corticais médias totais das duas classes.

Figura 40 – Distribuição da espessura cortical média total entre as classes CCLe e CCLp (com harmonização).



Fonte: Autoria própria.

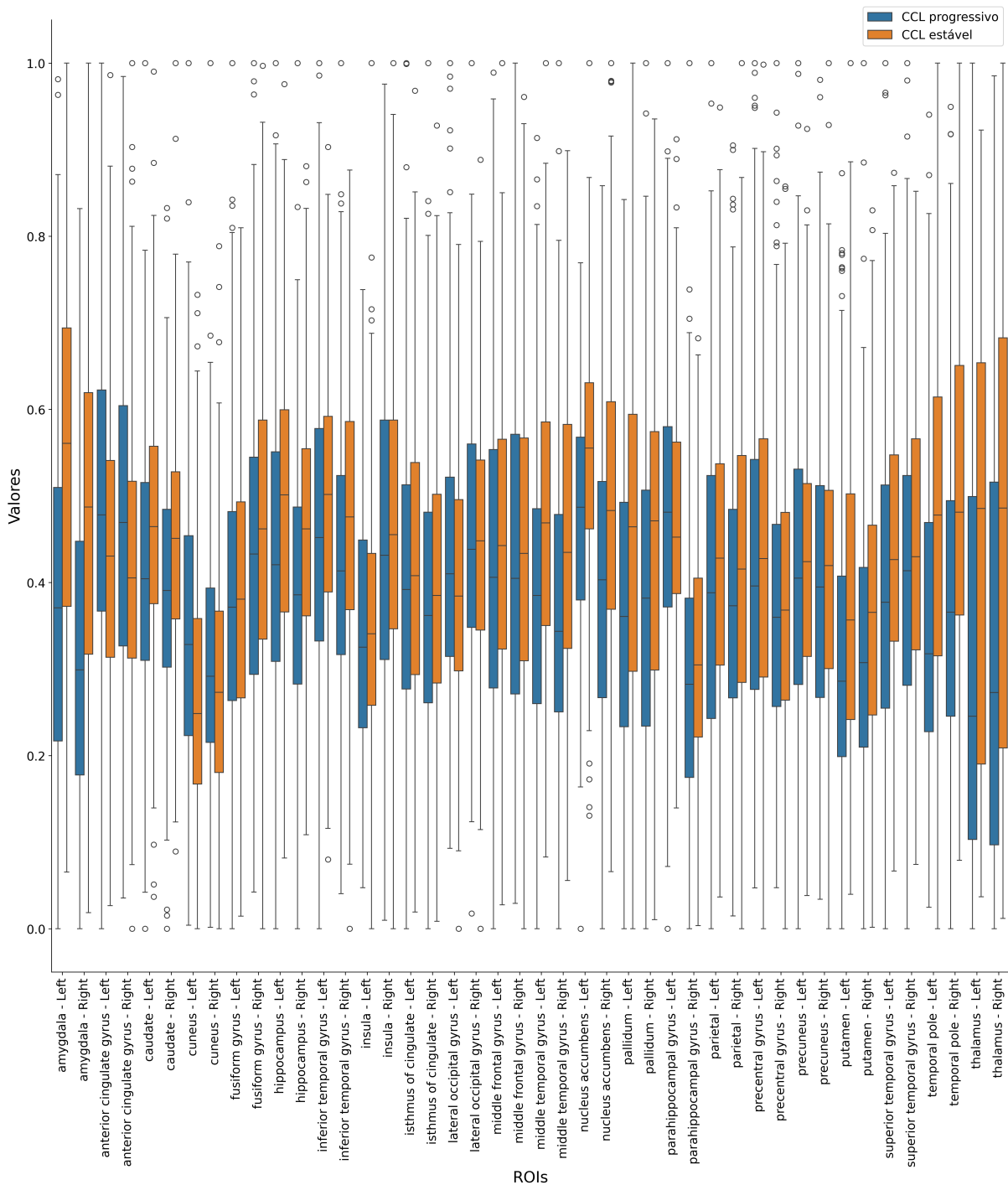
## 5.2.6 Espessura cortical por ROI

### 5.2.6.1 Sem harmonização ComBat

As espessuras corticais médias extraídas das 44 ROIs foram normalizados com o volume total intracraniano de cada paciente. A Figura 41 apresenta, para cada classe (CCLe e CCLp), os boxplots das medidas de espessura cortical média extraídas de cada ROI, com normalização por coluna para facilitar a visualização. Visualmente, observa-se que os boxplots da classe CCLe estão mais elevados, indicando uma tendência de valores mais altos para a espessura cortical média para essa classe. Além disso, a discrepância entre os boxplots das duas classes tem destaque para as regiões da amígdala, Área de Accumbens, pallidum, putâmen e tálamo.

Para uma análise quantitativa, foi realizado o teste de normalidade de Shapiro-Wilk, que indicou que 8 atributos seguem uma distribuição normal ( $p\text{-valor} > 0,05$ ). Em seguida, foi aplicado o teste de Levene para verificar a homogeneidade das variâncias, o qual confirmou a homoscedasticidade para 6 atributos e rejeitou para 2. Para verificar a significância das diferenças entre as classes CCLe e CCLp, utilizou-se o teste t de Student nos atributos com distribuição normal e homoscedasticidade, revelando que 5 deles apresentaram diferenças estatisticamente significativas. Para os atributos restantes, aplicou-se o

Figura 41 – Distribuição da espessura cortical média de cada ROI nas classes CCLe e CCLp (sem harmonização).



Fonte: Autoria própria.

teste de Welch, que identificou diferenças significativas em 24 atributos. Assim, obteve-se um total de 29 atributos com diferenças estatisticamente significativas entre as classes CCLe e CCLp.

### 5.2.6.2 Com harmonização ComBat

Para uma análise quantitativa dos atributos após a harmonização ComBat, foi realizado o teste de normalidade de Shapiro-Wilk, que indicou que 14 atributos seguem uma distribuição normal ( $p$ -valor  $> 0,05$ ). Em seguida, foi aplicado o teste de Levene para verificar a homogeneidade das variâncias, o qual confirmou a homoscedasticidade para 8 atributos e rejeitou para 6. Para verificar a significância das diferenças entre as classes CCLe e CCLp, utilizou-se o teste  $t$  de Student nos atributos com distribuição normal e homoscedasticidade, revelando que 5 deles apresentaram diferenças estatisticamente significativas. Para os atributos restantes, aplicou-se o teste de Welch, que identificou diferenças significativas em 13 atributos. Assim, obteve-se um total de 18 atributos com diferenças estatisticamente significativas entre as classes CCLe e CCLp.

A Figura 42 exibe os boxplots das medidas de espessura cortical média extraídas de cada ROI para cada classe, após a harmonização, com normalização por coluna para melhor visualização. Observa-se visualmente uma maior sobreposição dos boxplots entre as classes, o que pode explicar a redução na quantidade de atributos estatisticamente significativos após a harmonização. Após a harmonização, 18 atributos permaneceram significativos e 11 perderam a significância.

## 5.2.7 Atributos de textura da imagem toda

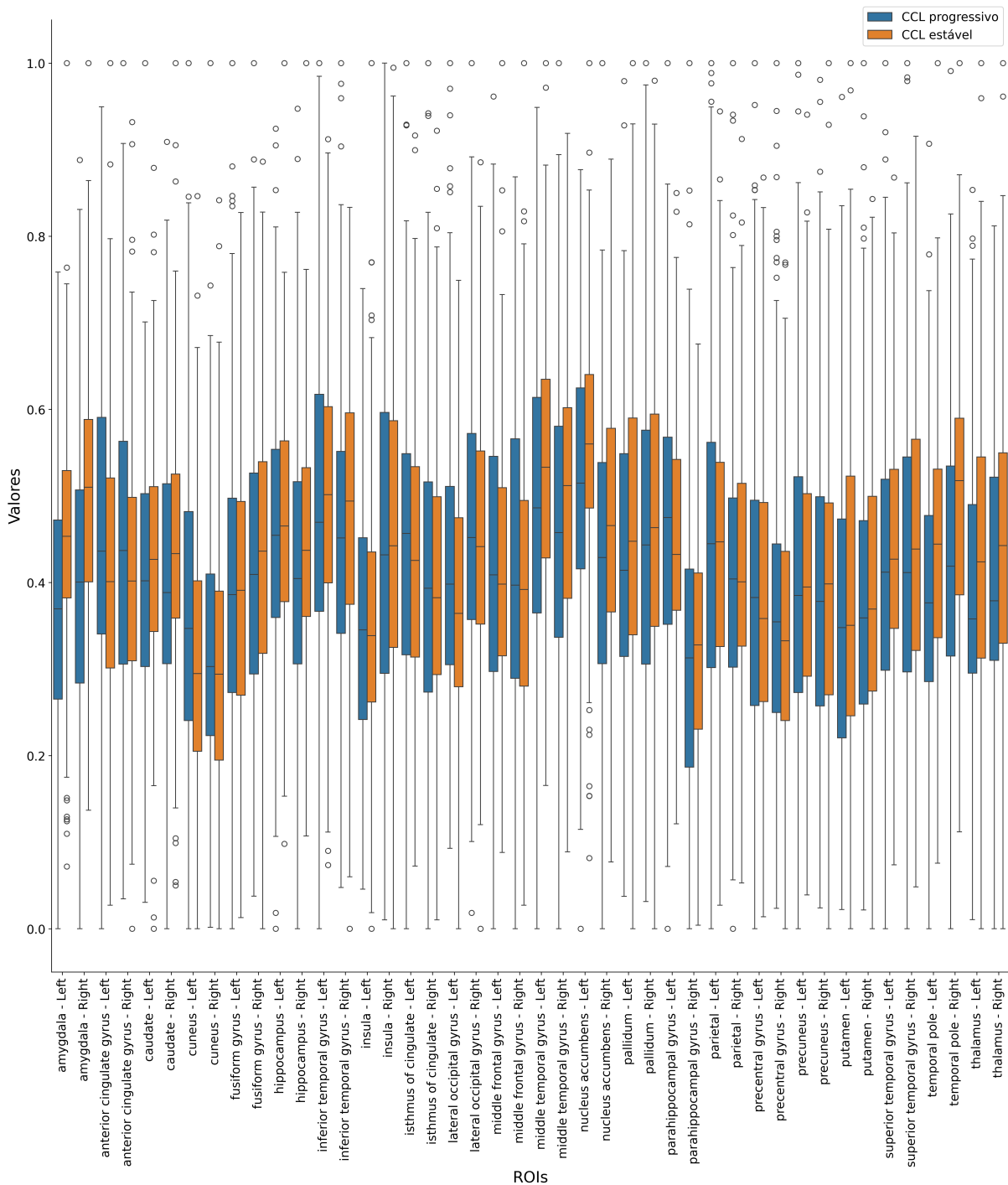
As medidas de textura podem ser influenciadas pelas etapas de pré-processamento das imagens. Diante disso, para fins de comparação, a extração dos atributos de textura foi realizada a partir de imagens com três tipos de tratamentos diferentes: imagens obtidas da etapa de remoção do crânio, ajustadas para resolução isotrópica de  $1 \text{ mm}^3$  e normalizada (imagens com ruído) (Tipo I); imagens pré-processadas (com remoção do crânio, redução de ruído e correção do *bias field*) e normalizadas (Tipo II); imagens pré-processadas e com padronização de histogramas (Tipo III). Além disso, para avaliar o impacto de diferentes níveis de quantização na construção das matrizes GLCM, os atributos de textura foram extraídos considerando quantizações de 32 e 64 bins.

Para avaliar estatisticamente os atributos extraídos, foi realizado o teste de normalidade de Shapiro-Wilk. Em seguida, o teste de Levene foi aplicado para verificar a homogeneidade das variâncias nos casos com distribuição normal. Para analisar a significância das diferenças entre as classes CCLe e CCLp, utilizou-se o teste  $t$  de Student para atributos com distribuição normal e homoscedasticidade, enquanto, para os demais, foi aplicado o teste de Welch.

### 5.2.7.1 Sem harmonização ComBat

Todos os atributos extraídos das imagens do Tipo I usando a GLCM de 32 bins tiveram sua normalidade rejeitada, e 13 deles apresentaram diferenças estatisticamente significa-

Figura 42 – Distribuição da espessura cortical média de cada ROI nas classes CCLe e CCLp (com harmonização).



Fonte: Autoria própria.

tivas entre as classes CCLp e CCLe. Para a GLCM de 64 bins, apenas a Correlação da Informação Mútua (*Mutual Information Correlation*, MIC ) I apresentou distribuição normal (mas teve sua homoscedasticidade rejeitada), enquanto os demais tiveram sua normalidade rejeitada. No total, 11 atributos mostraram diferenças estatisticamente significativas entre as classes.

Os 24 atributos extraídos das imagens do Tipo II usando a GLCM de 32 bins tiveram sua normalidade rejeitada, e 19 deles apresentaram diferenças estatisticamente significativas entre as classes CCLp e CCLe. Para a GLCM de 64 bins, apenas o atributo MIC I apresentou distribuição normal e homoscedasticidade, enquanto os demais tiveram sua normalidade rejeitada. No total, 21 atributos mostraram diferenças estatisticamente significativas entre as classes.

Dentre os atributos extraídos das imagens do Tipo III usando a GLCM de 32 bins, apenas o atributo MIC II apresentou distribuição normal e homoscedasticidade, enquanto os demais tiveram sua normalidade rejeitada. No total, apenas 7 atributos mostraram diferenças estatisticamente significativas entre as classes. Para a GLCM de 64 bins, apenas os atributos MIC I e correlação apresentaram distribuição normal e homoscedasticidade, enquanto os demais tiveram sua normalidade rejeitada. No total, apenas 8 atributos mostraram diferenças estatisticamente significativas entre as classes.

A Tabela 17 apresenta os atributos de textura extraídos da imagem completa do encéfalo que mostraram diferenças estatisticamente significativas entre CCLe e CCLp para cada tipo de imagem e nível de quantização, indicados pelo valor “X”. Quatro atributos foram significativos em todos os experimentos — MIC II, correlação, Proeminência do Agrupamento (*Cluster Prominence*) e Coeficiente Máximo de Correlação (MCC) —, destacando o potencial desses atributos para diferenciar as classes. Além disso, a extração dos atributos de textura da imagem toda do encéfalo que utilizou as imagens do Tipo II e quantização de 64 bins foi a que apresentou o maior número de atributos com diferenças estatisticamente significativas entre CCLe e CCLp. Isso sugere que essa configuração é a mais sensível para captar variações entre os grupos, podendo ser a mais adequada para discriminar entre essas classes.

#### 5.2.7.2 Com harmonização ComBat

Após a harmonização ComBat, todos os atributos extraídos das imagens do Tipo I usando a GLCM de 32 bins tiveram sua normalidade rejeitada, e nenhum deles apresentou diferenças estatisticamente significativas entre as classes CCLp e CCLe. O mesmo resultado foi observado para a GLCM de 64 bins.

Os 24 atributos extraídos das imagens do Tipo II usando a GLCM de 32 bins também tiveram sua normalidade rejeitada, e nenhum deles apresentaram diferenças estatisticamente significativas entre as classes CCLp e CCLe. Para a GLCM de 64 bins, apenas o atributo MIC I apresentou distribuição normal e homoscedasticidade, enquanto os demais tiveram sua normalidade rejeitada. Nenhum dos atributos mostraram diferenças estatisticamente significativas entre as classes.

Todos os atributos extraídos das imagens do Tipo III usando a GLCM de 32 bins tiveram sua normalidade rejeitada, e nenhum deles apresentou diferenças estatisticamente

Tabela 17 – Atributos de textura da imagem toda para cada tipo de imagem e nível de quantização que apresentaram diferenças estatisticamente significativas entre CCLe e CCLp (marcados com X) (sem harmonização).

Tipo das imagens de entrada	Tipo I (*)		Tipo II (**)		Tipo III (***)	
	32	64	32	64	32	64
<i>Cluster Shade</i>	X	-	X	X	-	-
Probabilidade Máxima	-	X	-	X	-	-
Média da Soma	-	-	-	-	-	-
MIC I	X	X	-	X	X	X
MIC II	X	X	X	X	X	X
Energia	X	X	X	X	-	-
Entropia	X	X	X	X	-	-
Entropia da Diferença	-	-	X	X	-	X
Correlação	X	X	X	X	X	X
Autocorrelação	-	-	-	-	-	-
Proeminência do Agrupamento ( <i>Cluster Prominence</i> )	X	X	X	X	X	X
Tendência do Agrupamento ( <i>Cluster Tendency</i> )	X	X	X	X	X	-
Contraste	-	-	X	X	-	X
Média da Diferença	-	-	X	X	-	-
Variância da Diferença	X	-	X	X	-	X
Diferença Inversa (ID)	-	-	X	X	-	-
Momento da Diferença Inversa (IDM)	-	-	X	X	-	-
Momento da Diferença Inversa Normalizado (IDMN)	X	-	X	X	-	-
Diferença Inversa Normalizada (IDN)	-	-	X	X	-	-
Variância Inversa	-	-	X	X	-	-
Média Conjunta	-	-	-	-	-	-
Coefficiente Máximo de Correlação (MCC)	X	X	X	X	X	X
Entropia da Soma	X	X	X	X	-	-
Quadrados da Soma	X	X	X	X	X	-
<b>Total</b>	<b>13</b>	<b>11</b>	<b>19</b>	<b>21</b>	<b>7</b>	<b>8</b>

(\*) Tipo I: imagens obtidas da etapa de remoção do crânio, ajustadas para resolução isotrópica de 1 mm e normalizadas;

(\*\*) Tipo II: imagens pré-processadas (com remoção do crânio, redução de ruído e correção do bias field) e normalizadas;

(\*\*\*) Tipo III: imagens pré-processadas e com padronização de histogramas.

Fonte: Autoria própria.

significativas entre as classes CCLp e CCLe. O mesmo resultado foi observado para a GLCM de 64 bins.

Diante disso, observa-se que, após o processo de harmonização ComBat, para todos os tipos de imagens testadas, nenhum atributo de textura para imagem toda apresentou diferenças estatisticamente significativas entre as classes CCLe e CCLp.

## 5.2.8 Atributos de textura de cada ROI

### 5.2.8.1 Sem harmonização ComBat

Entre os 1056 atributos extraídos das imagens do Tipo I utilizando a GLCM de 32 bins, apenas 86 apresentaram distribuição normal, dos quais 76 exibiram homoscedasticidade. No total, 642 atributos mostraram diferenças estatisticamente significativas entre as classes CCLp e CCLe. Resultados idênticos foram obtidos para a GLCM de 64 bins.

Para as imagens do Tipo II, 68 atributos extraídos com a GLCM de 32 bins apresentaram distribuição normal, dos quais 63 exibiram homoscedasticidade. No total, 708 atributos apresentaram diferenças estatisticamente significativas entre as classes, com os mesmos resultados observados para a GLCM de 64 bins.

Entre os atributos extraídos das imagens do Tipo III com a GLCM de 32 bins, 64 exibiram distribuição normal, dos quais 60 apresentaram homoscedasticidade. No total, 432 atributos mostraram diferenças estatisticamente significativas entre as classes, com resultados idênticos para a GLCM de 64 bins.

Não foram observadas diferenças entre as quantizações dentro de cada tipo de imagem, resultando em distribuições idênticas e no mesmo conjunto de atributos estatisticamente significativos. Observou-se que 245 atributos de textura extraídos das ROIs estavam presentes nos conjuntos de atributos estatisticamente significativos em todos os tipos de imagens analisados. Esses atributos correspondem a 33 ROIs, conforme apresentado na Tabela 18, abrangendo um ou ambos os hemisférios de 20 das 22 regiões selecionadas neste estudo, com exceção do Precuneus e do Pallidum. Além disso, a única medida de textura ausente foi a 'Idmn'.

Entre os conjuntos de atributos com diferenças estatisticamente significativas entre as classes, o Tipo II foi o único a incluir a região Área de Accumbens do hemisfério esquerdo, ausente nos demais conjuntos.

Tabela 18 – ROIs que tiveram algum atributo de textura com diferença estatisticamente significativa em todos os tipos de imagens testados (sem harmonização).

ID	Região	ID	Região
1	Hipocampo (direito e esquerdo)	11	Polo temporal (direito e esquerdo)
2	Amígdala (direito e esquerdo)	12	Giro fusiforme (direito e esquerdo)
3	Giro temporal médio (direito e esquerdo)	13	Parietal (direito e esquerdo)
4	Giro parahipocampal (direito e esquerdo)	14	Giro occipital lateral (direito e esquerdo)
5	Giro temporal inferior (direito e esquerdo)	15	Cuneus (direito)
6	Giro temporal superior (direito e esquerdo)	16	Giro pré-central (direito)
7	Caudado (direito)	17	Istmo do cíngulo (esquerdo)
8	Ínsula (direito e esquerdo)	18	Giro frontal médio (direito e esquerdo)
9	Putâmen (direito)	19	Giro cingulado anterior (esquerdo)
10	Área de Accumbens (direito)	20	Tálamo (direito e esquerdo)

Fonte: Autoria própria.

### 5.2.8.2 Com harmonização ComBat

Após a harmonização ComBat, entre os 1056 atributos extraídos das imagens do Tipo I utilizando a GLCM de 32 bins, apenas 167 apresentaram distribuição normal, dos quais 156 exibiram homoscedasticidade. No total, 269 atributos mostraram diferenças estatisticamente significativas entre as classes CCLp e CCLe. Para a GLCM de 64 bins, as mesmas distribuições e homoscedasticidades foram observadas, porém, apenas 267 atributos mostraram diferenças estatisticamente significativas entre as classes. A quantização de 32 bins incluiu todos os atributos significantes obtidos pela GLCM de 64 bins e incluiu dois novos atributos (Cluster Prominence do Giro Temporal Inferior esquerdo e Probabilidade Máxima do Istmo do cíngulo direito).

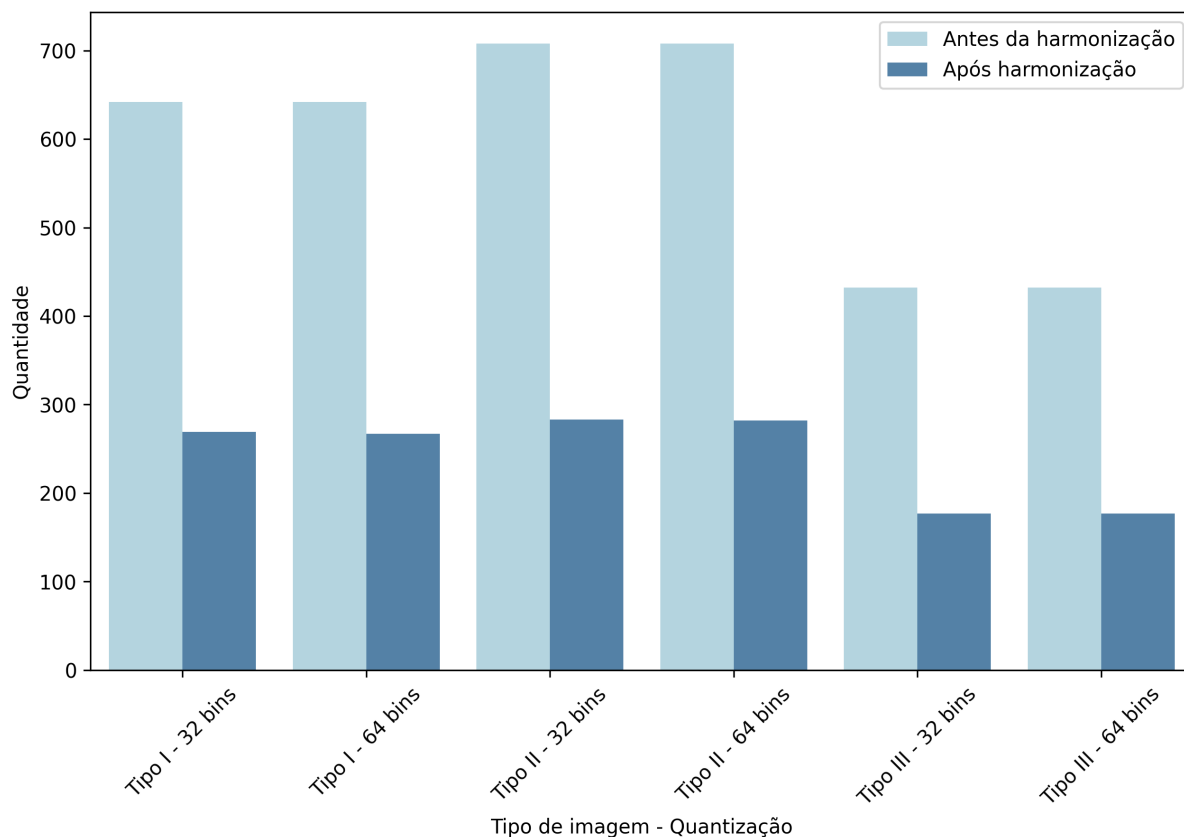
Para as imagens do Tipo II, 151 atributos extraídos com a GLCM de 32 bins apresentaram distribuição normal, dos quais 148 exibiram homoscedasticidade. No total, 283 atributos apresentaram diferenças estatisticamente significativas entre as classes. As mesmas distribuições e homoscedasticidades foram observadas para a GLCM de 64 bins, porém, para essa quantização apenas 282 atributos apresentaram diferenças estatisticamente significativas entre as classes. A quantização de 32 bins incluiu todos os atributos significantes obtidos pela GLCM de 64 bins e incluiu um novo atributo (Soma da Entropia do Tálamo esquerdo).

Entre os atributos extraídos das imagens do Tipo III com a GLCM de 32 bins, 80 exibiram distribuição normal, os quais também apresentaram homoscedasticidade. No total, 177 atributos mostraram diferenças estatisticamente significativas entre as classes. Para a GLCM de 64 bins, 81 exibiram distribuição normal, os quais também apresentaram homoscedasticidade. Entretanto, também apresentou 177 atributos que mostraram diferenças estatisticamente significativas entre as classes. A quantização de 32 bins incluiu todos os atributos significantes obtidos pela GLCM de 64 bins, exceto o Cluster Prominence do Giro Médio esquerdo, e incluiu um novo atributo (Correlação da amígdala

esquerda).

Diante disso, conforme apresentado na Figura 43, após a harmonização ComBat, houve diferenças nos resultados ao usar diferentes quantizações para cada tipo de imagem testado. Além disso, é possível visualizar que o número de atributos estatisticamente significativos diminuiu para todos os experimentos após a harmonização ComBat.

Figura 43 – Quantidade de atributos estatisticamente significativos para cada tipo de imagem e quantização antes e após harmonização ComBat.



Fonte: Autoria própria.

Além disso, verificou-se que, após a harmonização, enquanto alguns atributos permaneceram estatisticamente significativos, outros perderam essa significância, e alguns passaram a ser significativos.

No caso das imagens do Tipo I, os mesmos 31 atributos tornaram-se significativos para a GLCM de 32 e 64 bins, enquanto 404 e 406 atributos perderam a significância, respectivamente, para cada quantização. Para as imagens do Tipo II, os mesmos 25 atributos ganharam significância para a GLCM de 32 e 64 bins, enquanto 450 e 451 atributos perderam a significância, respectivamente, para cada quantização.

Por fim, para as imagens do Tipo III, tanto na GLCM de 32 bins quanto na de 64 bins, 61 atributos passaram a ser significativos, enquanto 315 atributos em comum perderam a significância, além de um atributo distinto para cada quantização que também perdeu significância.

Diante dos resultados, os atributos de textura extraídos de imagens do Tipo II utilizando GLCM com 64 bins foram selecionados para compor o conjunto de atributos utilizados no treinamento dos modelos de classificação propostos. Essa escolha foi motivada pelo maior número de atributos com diferença estatisticamente significativa entre as classes CCLe e CCLp quando as imagens do Tipo II foram empregadas.

Quanto à quantização, as configurações com 32 e 64 bins apresentaram os mesmos resultados para os atributos de textura extraídos da imagem inteira (com harmonização) e para os atributos de textura das ROIs (sem harmonização). No entanto, para os atributos de textura da imagem inteira sem harmonização, a GLCM com 64 bins não apenas incluiu todos os atributos significativos da GLCM com 32 bins, mas também identificou dois atributos adicionais, enquanto a GLCM com 32 bins apresentou apenas um atributo significativo adicional nas texturas de ROIs sem harmonização.

### 5.3 Harmonização ComBat

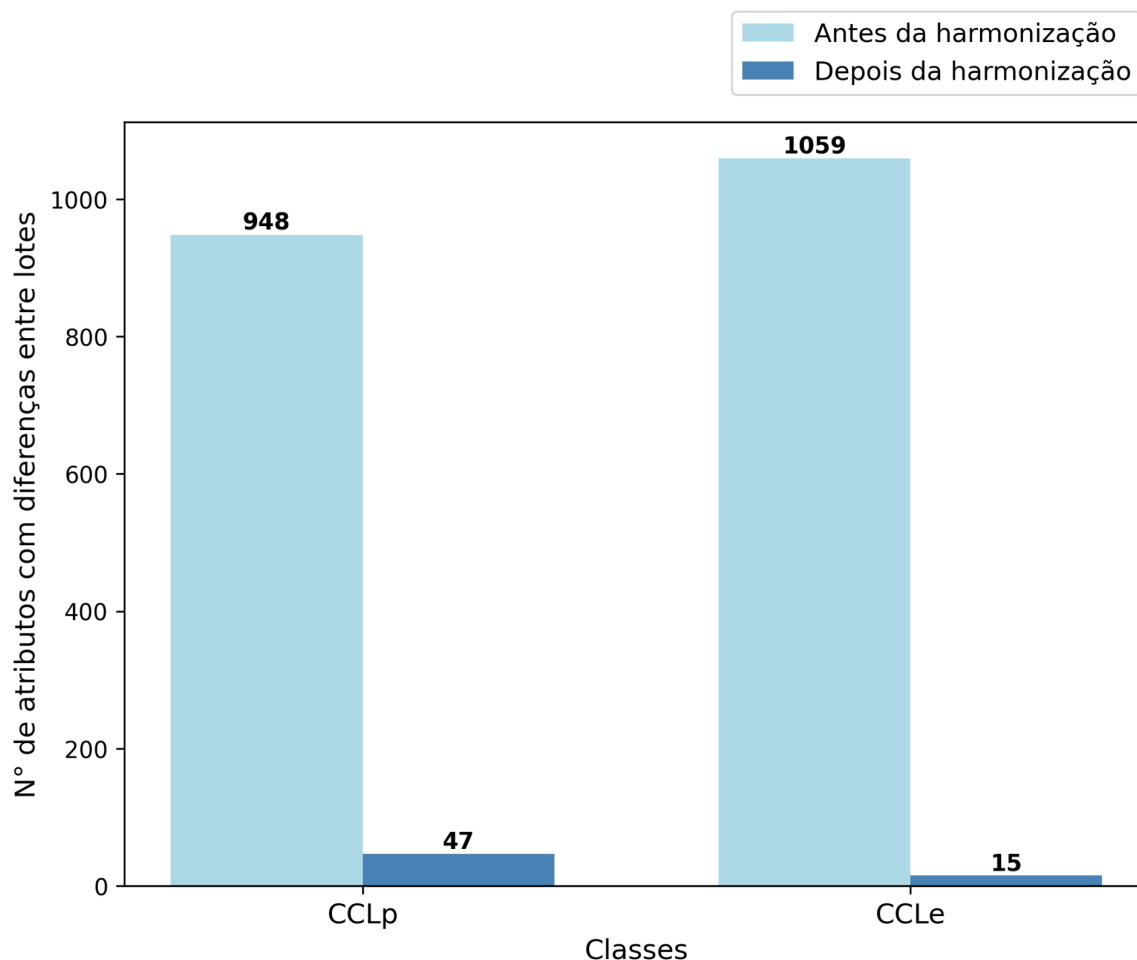
Para aplicar a harmonização pelo método ComBat, foi necessário, inicialmente, definir o conjunto de atributos a ser harmonizado, incluindo todos os tipos de atributos extraídos. Para a composição dos atributos de textura, foram selecionados aqueles obtidos a partir das imagens do Tipo II utilizando GLCM com 64 bins. A harmonização dos atributos extraídos de outras configurações de tipos de imagens e quantizações foi realizada seguindo o mesmo procedimento descrito neste trabalho, com a substituição dos conjuntos de atributos de textura conforme cada configuração.

Em seguida, após definir cada intensidade de campo (1,5 T e 3T) como sendo um lote, foi aplicado o teste de normalidade de Shapiro-Wilk para verificar a distribuição de cada atributo em cada lote. Com isso, verificou-se que 86% dos atributos do lote de 1,5 T e 78% de 3T não seguem distribuição normal. Com isso, definiu-se que o ComBat não paramétrico seria o mais apropriado.

Antes da aplicação da harmonização, conforme a Figura 44, verificou-se que 1059 atributos da classe CCLe e 948 atributos de CCLp apresentavam diferenças estatisticamente significativas entre os lotes. Após a harmonização ComBat não paramétrica, essas quantidades reduziram para 15 na classe CCLe e 47 na classe CCLp, apesar de a classe CCLp inicialmente apresentar um número menor de atributos com diferenças significativas.

A Figura 45 e a Figura 46 mostram o efeito da harmonização ComBat sobre cada conjunto de atributos das classes CCLe e CCLp, respectivamente. Observa-se que, em ambas as classes, os atributos de textura extraídos das ROIs foram os que mais apresentaram diferenças estatisticamente significativas entre os lotes, sugerindo que esses atributos são sensíveis às variações de intensidade de campo dos *scanners*. Além disso, os mesmos conjuntos de atributos exibiram quantidades muito próximas ao comparar as duas classes. O volume total intracraniano não demonstrou sensibilidade aos efeitos de lote. Em ambas

Figura 44 – Efeito da harmonização ComBat sobre os atributos das classes CCLe e CCLp.



Fonte: Autoria própria.

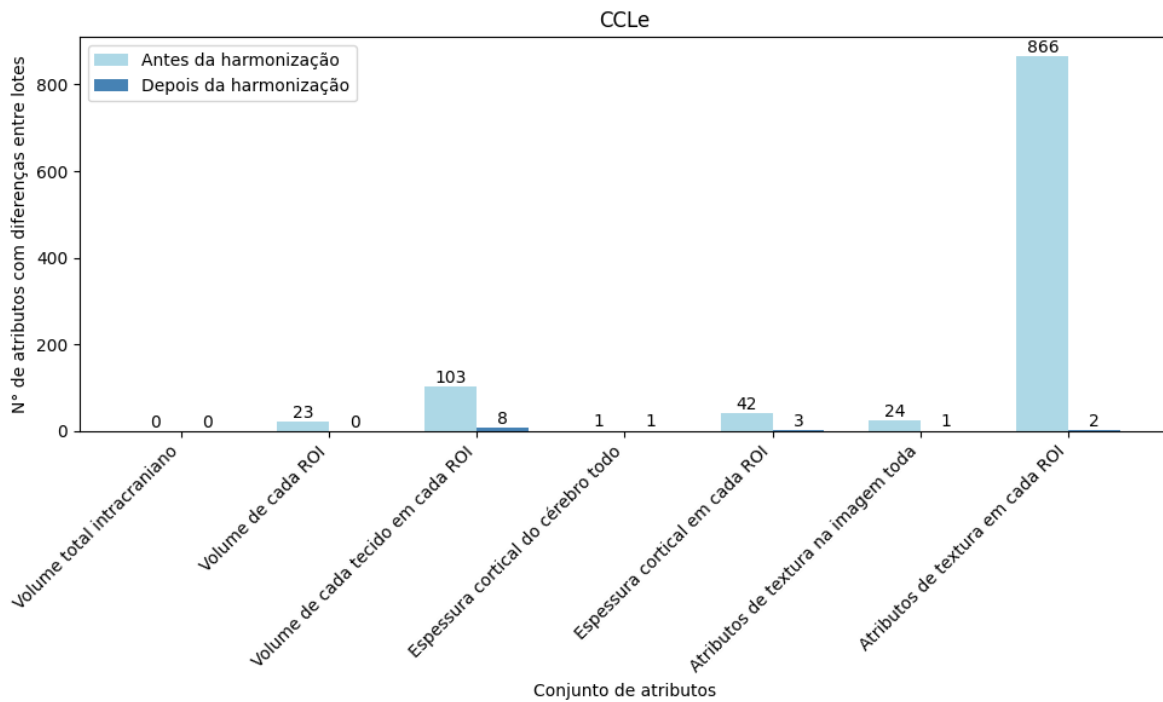
as classes, as diferenças entre lotes para os atributos de volume de cada ROI foram eliminadas após a harmonização. No entanto, no caso da classe CCLp, os efeitos de lote foram corrigidos em um número maior de conjuntos de atributos em comparação com a classe CCLe, incluindo a eliminação dos efeitos nos atributos de espessura cortical (cérebro todo e em cada ROI) e nos atributos de textura extraídos da imagem como um todo.

## 5.4 Experimentos de classificação

No total, foram treinados e testados, separadamente, 21 modelos de classificação baseados em Floresta Aleatória, considerando diferentes combinações de tratamentos aplicados aos dados. O procedimento de harmonização ComBat foi analisado sob três configurações: (i) não aplicado, (ii) aplicado a todos os atributos e (iii) aplicado com remoção subsequente dos atributos que ainda apresentavam efeitos de lote residuais.

As abordagens de seleção de atributos incluíram: sem seleção, agrupamento hierár-

Figura 45 – Efeito da harmonização ComBat sobre os conjuntos de atributos da classe CCLe.

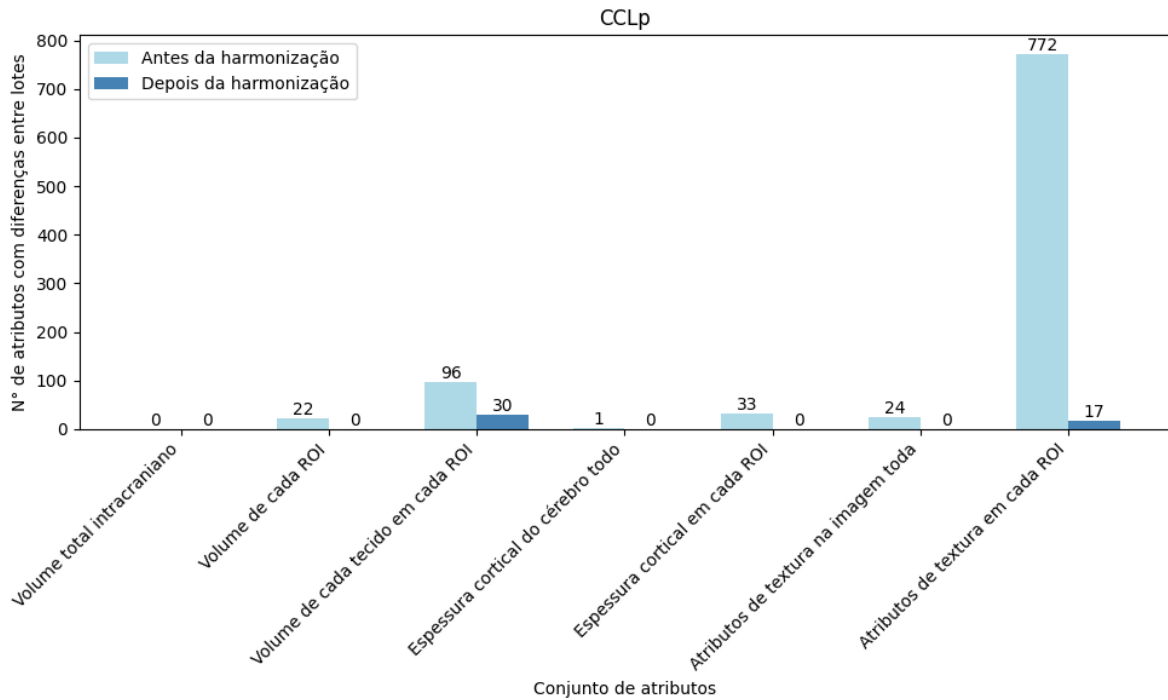


Fonte: Autoria própria.

quico e ranqueamento informado pelo modelo, com base nas pontuações de importância das características da Floresta Aleatória, com limiares de percentis variando de 10% a 90%.

Para os experimentos de classificação, o treinamento e a avaliação dos modelos foi conduzida por meio de NCV estratificado de 10 partes, combinada com uma NCV interna de 5 partes para ajuste de parâmetros. O melhor conjunto de hiperparâmetros foi selecionado com base no F1-score ponderado, que ajusta o F1-score para atribuir mais peso às classes com maior presença no conjunto de dados, levando em conta a média harmônica entre precisão e *recall*. O desempenho dos modelos foi avaliado usando as métricas de acurácia, acurácia balanceada, F1-score, sensibilidade, especificidade e AUC. A cada iteração do *loop* externo, o modelo ajustado foi avaliado nos dados de teste, e o modelo com a maior F1-score ponderado foi selecionado como o melhor classificador. A configuração de cada experimento, juntamente com seus melhores resultados e demais detalhes, estão resumidos na Tabela 19.

Figura 46 – Efeito da harmonização ComBat sobre os conjuntos de atributos da classe CCLp.



Fonte: Autoria própria.

Tabela 19 – Desempenho dos modelos de Floresta Aleatória sob diferentes configurações de pré-processamento de dados.

ID	Harmonização ComBat	Seleção de atributos	Percentil	Nº Atributos	Acc. (%)	Acc. Bal.(%)	Sens. (%)	Espec. (%)	AUC	F1-score (%)
1	Não aplicado	Não aplicado	-	1303	74,51 ± 4,12	74,31 ± 4,14	64,00 ± 7,57	84,62 ± 8,43	0,84 ± 0,06	74,21 ± 4,08
2	Não aplicado	Importância do RF	10	1172	72,55 ± 5,64	72,69 ± 5,59	75,00 ± 7,10	70,37 ± 10,42	0,74 ± 0,08	72,57 ± 5,76
3	Não aplicado	Importância do RF	25	977	74,51 ± 5,26	74,31 ± 5,18	64,00 ± 6,19	84,62 ± 10,36	0,84 ± 0,06	74,21 ± 5,27
4	Não aplicado	Importância do RF	50	652	78,43 ± 3,93	78,38 ± 3,84	76,00 ± 4,36	80,77 ± 8,27	0,83 ± 0,06	78,41 ± 3,95
5	Não aplicado	Importância do RF	75	326	76,47 ± 6,31	76,46 ± 6,39	76,00 ± 12,18	76,92 ± 9,87	0,83 ± 0,07	76,47 ± 6,46
6	Não aplicado	Importância do RF	90	131	76,47 ± 5,64	76,38 ± 5,58	72,00 ± 7,84	80,77 ± 10,51	0,84 ± 0,06	76,42 ± 5,65
7	Não aplicado	Clustering	-	956	74,51 ± 5,14	74,31 ± 5,08	64,00 ± 8,00	84,62 ± 12,01	0,82 ± 0,06	74,21 ± 5,26
8	Aplicado	Não aplicado	-	1303	74,51 ± 8,66	74,31 ± 8,62	64,00 ± 8,10	84,62 ± 13,19	0,86 ± 0,09	74,21 ± 8,65
9	Aplicado	Importância do RF	10	1172	72,55 ± 6,40	72,38 ± 6,47	64,00 ± 10,37	80,77 ± 11,08	0,84 ± 0,08	72,34 ± 6,44
10	Aplicado	Importância do RF	25	977	72,55 ± 5,99	72,46 ± 5,91	68,00 ± 5,99	76,92 ± 10,82	0,85 ± 0,08	72,49 ± 5,98
11	Aplicado	Importância do RF	50	652	78,43 ± 7,30	78,38 ± 7,25	76,00 ± 8,34	80,77 ± 12,26	0,83 ± 0,08	78,41 ± 7,31
12	Aplicado	Importância do RF	75	326	76,47 ± 6,22	76,38 ± 6,22	72,00 ± 10,10	80,77 ± 12,42	0,85 ± 0,07	76,42 ± 6,27
13	Aplicado	Importância do RF	90	131	76,47 ± 6,45	76,23 ± 6,34	64,00 ± 5,72	88,46 ± 12,72	0,88 ± 0,07	76,08 ± 6,52
14	Aplicado	Clustering	-	965	72,55 ± 6,05	72,46 ± 6,08	68,00 ± 9,85	76,92 ± 11,69	0,83 ± 0,08	72,49 ± 5,98
15	Remoção de atributos	Não aplicado	-	1246	76,47 ± 7,82	76,46 ± 7,76	76,00 ± 8,87	76,92 ± 11,39	0,81 ± 0,09	76,47 ± 7,94
16	Remoção de atributos	Importância do RF	10	1121	76,47 ± 7,57	76,38 ± 7,64	72,00 ± 11,12	80,77 ± 11,87	0,83 ± 0,08	76,42 ± 7,64
17	Remoção de atributos	Importância do RF	25	934	72,55 ± 7,19	72,46 ± 7,36	68,00 ± 12,46	76,92 ± 10,18	0,82 ± 0,08	72,49 ± 7,34
18	Remoção de atributos	Importância do RF	50	623	74,51 ± 7,27	74,38 ± 7,15	68,00 ± 9,28	80,77 ± 14,67	0,84 ± 0,07	74,39 ± 7,40
19	Remoção de atributos	Importância do RF	75	312	84,31 ± 8,70	84,31 ± 8,64	84,00 ± 9,67	84,62 ± 13,01	0,85 ± 0,07	84,31 ± 8,73
20	Remoção de atributos	Importância do RF	90	125	76,47 ± 7,38	76,38 ± 7,31	72,00 ± 8,16	80,77 ± 12,64	0,86 ± 0,08	76,42 ± 7,42
21	Remoção de atributos	Clustering	-	905	68,63 ± 4,69	68,54 ± 4,62	64,00 ± 9,16	73,08 ± 10,78	0,78 ± 0,07	68,55 ± 4,64

Fonte: Autoria própria.

Segundo esse critério, o Experimento 19 - o qual utilizou dados harmonizados com remoção de atributos que ainda apresentaram diferenças entre lotes pós-harmonização, e com 312 atributos selecionados a partir da importância de atributos na Floresta Aleatória - foi o que demonstrou maior F1-score (F1-score: 84,31±8,73%, ACC: 84,31±8,70%,

ACC balanceada:  $84,31 \pm 8,64\%$ , Sens.:  $84,00 \pm 9,67\%$ , Espec.:  $84,62 \pm 13,01\%$  e AUC:  $0,85 \pm 0,07$ ) e o experimento 21 - o qual utilizou dados harmonizados com remoção de atributos que ainda apresentaram diferenças entre lotes pós-harmonização, e com 905 atributos selecionados a partir da técnica de *clustering* - apresentou a menor F1-score (F1-score:  $68,55 \pm 4,64\%$ , ACC:  $68,63 \pm 4,69\%$ , ACC balanceada:  $68,54 \pm 4,62\%$ , Sens.:  $64,00 \pm 9,16\%$ , Espec.:  $73,08 \pm 10,78\%$  e AUC:  $0,78 \pm 0,07$ ).

Entre as cinco principais configurações, 3 foram treinadas com dados harmonizados e submetidos à seleção de atributos por meio da técnica de Importância da Floresta Aleatória. Além disso, na etapa de seleção de atributos, dois modelos utilizaram dados filtrados pelo percentil 50 e dois pelo percentil 75, enquanto um apenas removeu atributos que ainda apresentaram diferenças entre lotes pós-harmonização não utilizando um método extra de seleção de atributos.

Esses resultados sugerem que a harmonização dos dados e a seleção de atributos com base na Importância da Floresta Aleatória contribuem de maneira significativa para a qualidade dos modelos. Ainda que a métrica principal de avaliação tenha sido o F1-score, a análise da AUC evidenciou que os modelos que combinaram essas estratégias apresentaram maior capacidade discriminativa. Isso indica que tais abordagens não apenas favorecem o desempenho em um ponto específico de decisão, mas também promovem maior robustez ao longo de diferentes limiares, o que pode ser particularmente relevante em cenários clínicos.

Além disso, é possível observar na Figura 47 que os 10 piores modelos foram treinados com mais de 900 atributos, enquanto entre os 10 melhores modelos, apenas 2 utilizaram uma quantidade mais que esse valor. Isso sugere que, na maioria dos experimentos, um número elevado de atributos pode ter dificultado o ajuste adequado dos parâmetros, aumentando a complexidade do modelo e potencialmente levando ao *overfitting*.

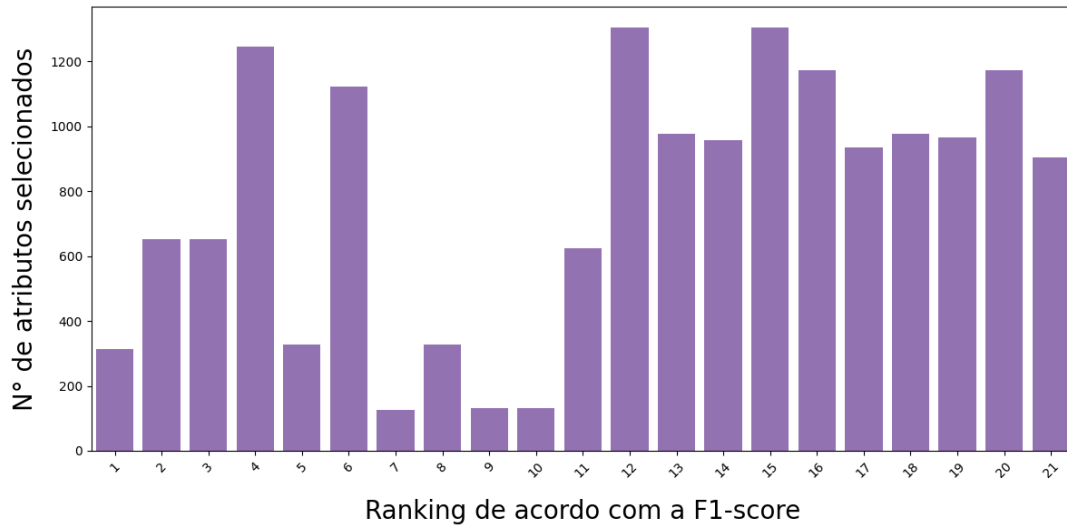
Outro ponto relevante é que modelos com menos atributos tendem a ter melhor desempenho, o que pode estar relacionado a uma melhor generalização dos padrões nos dados. Isso reforça a importância de métodos de seleção de atributos para eliminar variáveis redundantes ou pouco informativas, reduzindo o risco de ajuste excessivo ao conjunto de treinamento.

Para validar ainda mais a superioridade do modelo selecionado, trabalhos futuros podem incluir testes de significância estatística entre as dobras da validação cruzada. Testes pareados não paramétricos, como o teste de postos sinalizados de Wilcoxon, podem ser utilizados para comparar as distribuições de F1-score e AUC entre os modelos.

#### 5.4.0.1 Análise dos atributos - Experimento 19

O modelo do Experimento 19 apresentou o melhor desempenho segundo a F1-score. Esse resultado foi obtido com a aplicação da seleção de atributos baseada na importância

Figura 47 – Relação entre F1-score e quantidade de atributos selecionados.



Fonte: Autoria própria.

atribuída pela Floresta Aleatória, considerando o percentil 75. Foram selecionados 312 atributos, abrangendo as 66 ROIs (descritas na Subseção 4.5).

Para analisar os atributos mais relevantes para a previsão do modelo, utilizou-se a abordagem SHAP (vide Subseção 4.8.4), analisando cada classe separadamente. A Figura 48 exibe os 20 atributos mais importantes para a previsão de CCL progressivo no Experimento 19, destacando tanto sua importância global quanto suas respectivas contribuições locais. Descrições detalhadas dos atributos apresentados nessa figura estão disponíveis no Apêndice 7.2.

Observa-se que 11 dos 20 atributos mais relevantes estão relacionados ao volume (5 de GM, 5 de WM e 1 volume total da ROI giro temporal inferior direito), 6 à textura (2 referentes ao IDMN, 1 ao IDN, 1 à Média da Soma, 1 à Energia e 1 à Média da Diferença) e 3 à espessura cortical. Além disso, 9 dos 20 atributos estão associados à amígdala, 6 a regiões de giros cerebrais, 2 à Área de Accumbens, 2 ao polo temporal e 1 ao hipocampo. No total, os 20 atributos mais relevantes estão distribuídos entre 8 regiões cerebrais, sendo 8 no hemisfério esquerdo e 12 no hemisfério direito.

Entre os atributos volumétricos mais importantes, destacam-se a amígdala (GM e WM, em ambos os hemisférios), o giro temporal inferior (volume total e GM no hemisfério direito), o giro temporal médio (GM em ambos os hemisférios), a Área de Accumbens (WM no hemisfério esquerdo), giro fusiforme (WM no hemisfério direito) e polo temporal (WM no hemisfério esquerdo). Quanto aos atributos de espessura cortical, as regiões mais relevantes incluem a amígdala (esquerda e direita) e o polo temporal direito. Já entre os atributos de textura, destacam-se aqueles relacionados ao hipocampo direito (como o IDMN), à amígdala direita (IDN, IDMN e Energia), à Área de Accumbens esquerda (Média da Soma) e ao giro temporal superior esquerdo (Média da Diferença).

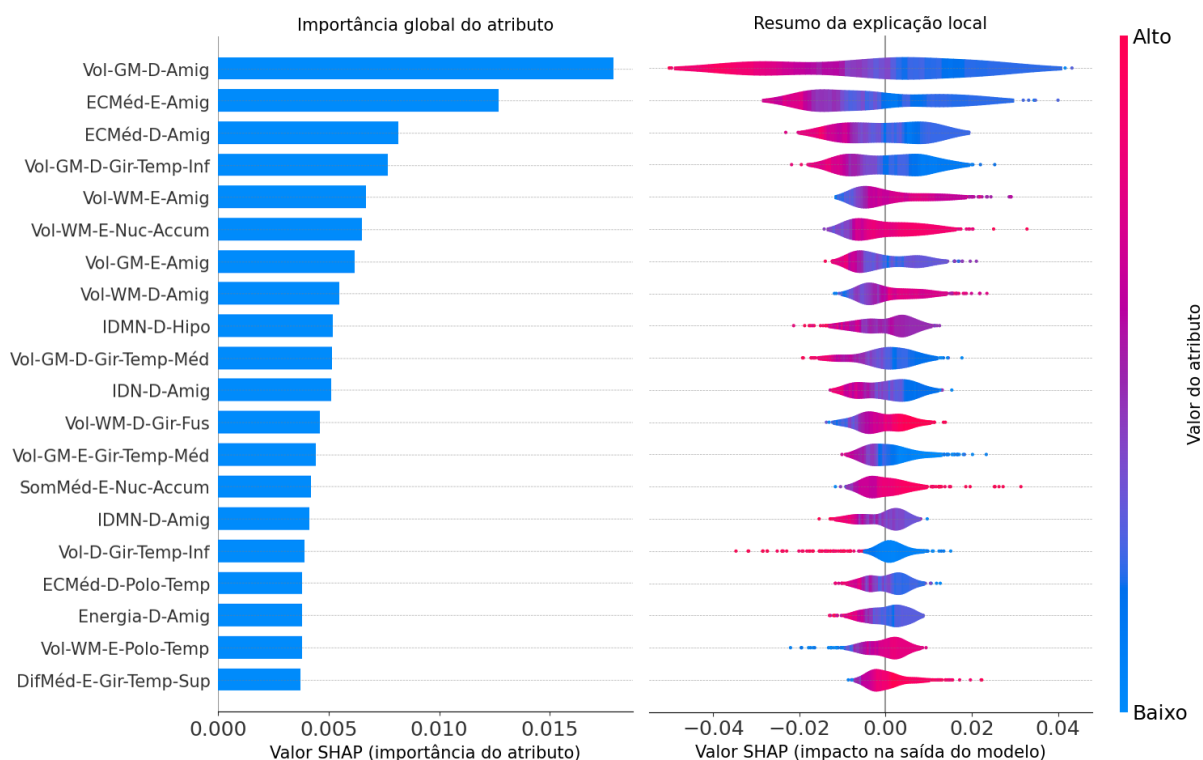
Entre os atributos de textura identificados como relevantes para a classificação de CCLp, observou-se que os valores SHAP associados ao IDMN, IDN e à Energia foram predominantemente negativos para valores altos desses atributos. Isso indica que texturas mais homogêneas e regulares, especialmente no hipocampo direito e na amígdala direita, estão associadas a uma menor probabilidade de conversão, refletindo possivelmente preservação microestrutural dessas regiões. Por outro lado, os atributos Sum Average (área de accumbens esquerda) e Difference Average (giro temporal superior esquerdo) apresentaram valores SHAP positivos para valores elevados, sugerindo que a presença de intensidades mais altas ou maior variação local nos níveis de cinza podem estar relacionadas a alterações texturais compatíveis com neurodegeneração precoce. Esses achados reforçam a relevância de padrões de textura como possíveis biomarcadores não invasivos na detecção de alterações estruturais sutis associadas à progressão da doença.

O volume de GM da amígdala direita foi o atributo de maior importância global. De acordo com sua explicação local, valores elevados desse atributo reduzem a probabilidade da instância ser classificada como CCLp. O mesmo comportamento foi observado para a espessura cortical média da amígdala esquerda, o segundo atributo mais importante. Em contrapartida, volumes de WM (da amígdala esquerda e direita, da Área de Accumbens esquerda, giro fusiforme direito e polo temporal esquerdo), a Média da Soma da Área de Accumbens esquerda e a Média da Diferença do giro temporal superior esquerdo, apresentaram um efeito oposto, onde valores mais altos aumentam a probabilidade de classificação como CCLp.

A Figura 49 sugere que há uma relação inversa entre os volumes de GM e WM na amígdala esquerda, corroborando a hipótese de que a progressão para CCLp está associada a uma redução do volume de GM e um aumento compensatório de WM. Descrições detalhadas dos atributos apresentados nessa figura estão disponíveis no Apêndice 7.2. Essa relação é evidenciada pelos valores SHAP positivos para baixos volumes de GM, associados a volumes maiores de WM (em vermelho). Por outro lado, volumes maiores de GM, frequentemente associados a volumes menores de WM (em azul), enfraquecem a predição da classe CCLp, conforme indicado pelos valores SHAP negativos, indicando que essa característica contribui negativamente para a predição de CCLp.

Esse padrão volumétrico inverso GM e WM pode refletir uma combinação de atrofia neurodegenerativa na GM — como perda neuronal e regressão dendrítica — e glioses reativas ou efeitos de volume parcial na substância branca adjacente, que podem levar a aumentos aparentes no volume de WM em imagens de ressonância magnética estrutural. Adicionalmente, a desmielinização precoce e a reclassificação induzida pela segmentação na fronteira entre GM e WM podem acentuar ainda mais essa tendência inversa.

Figura 48 – Importância global e contribuições locais dos 20 atributos mais importantes do modelo do Experimento 19 para previsão de CCLp.



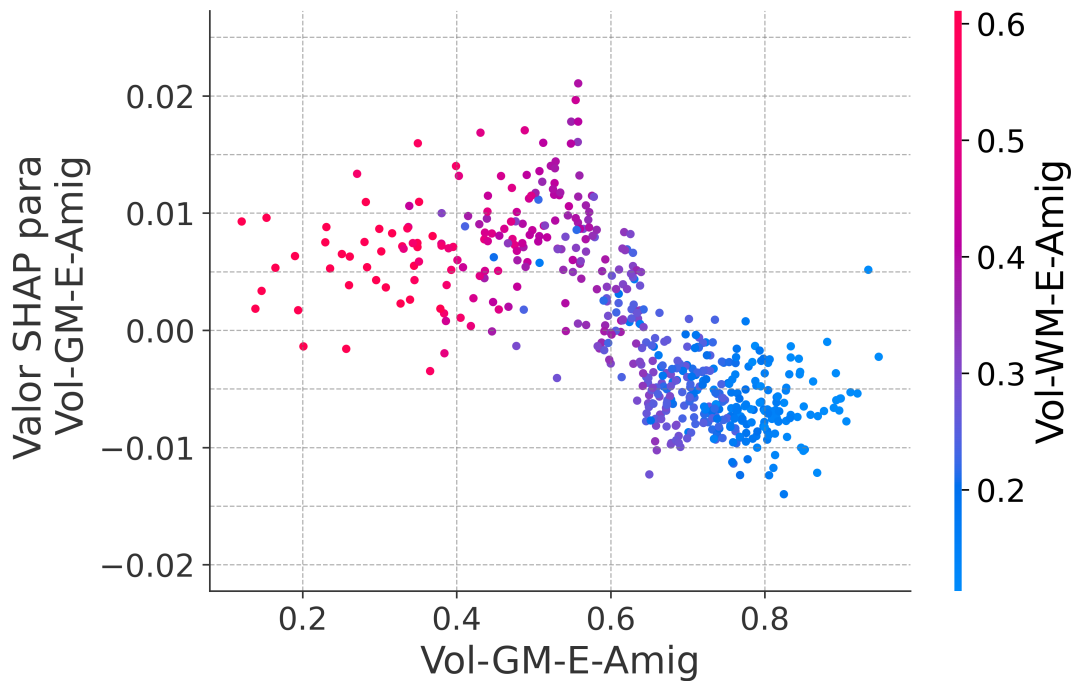
Fonte: Autoria própria.

## 5.5 Comparação com trabalhos da literatura

Os resultados obtidos neste trabalho mostraram-se comparáveis aos reportados na literatura. Considerando os estudos baseados em aprendizagem profunda, o F1-score do Experimento 19 ( $84,31 \pm 8,73\%$ ) foi inferior apenas ao trabalho de Ashtari-Majlan, Seifi e Dehshibi (2022) (F1-score: 85,96%), mas superou o resultado reportado por Luo et al. (2023) (F1-score: 75%). Em comparação com estudos baseados em engenharia de atributos, o F1-score do Experimento 19 foi inferior aos alcançados por Syaifullah et al. (2021), que reportaram F1-score de 86,9% para o modelo SVMest e 88,5% para o modelo SVMcog. Vale destacar que, entre todos os estudos levantados nesta revisão, apenas esses quatro trabalhos reportaram explicitamente a métrica F1-score, o que limita comparações diretas, por essa métrica, com outros métodos da literatura.

Considerando os estudos baseados em aprendizagem profunda, a AUC do Experimento 19 ( $AUC: 0,85 \pm 0,07$ ) foi superada apenas pelos trabalhos de Ashtari-Majlan, Seifi e Dehshibi (2022) ( $AUC: 0,94$ ) e Rahim et al. (2025) ( $AUC: 0,89$ ). Quando comparado aos estudos baseados em engenharia de atributos, a AUC do Experimento 19 foi inferior aos de Minhas et al. (2021) ( $AUC: 0,92$ ), Syaifullah et al. (2021) ( $AUC: 0,91$ ) e Xie et al. (2023) ( $AUC: 0,88$ ), porém superou os resultados de seis outros trabalhos da literatura (LEE; LEE; KIM, 2020; GAO et al., 2020; BAE et al., 2021; SHU et al., 2021;

Figura 49 – Interação entre os volumes de GM e WM da amígdala esquerda na predição de CCLp pelo modelo do Experimento 19.



Fonte: Autoria própria.

CALANDRELLI et al., 2022; LUO et al., 2023).

Em relação à acurácia do Experimento 19 (ACC:  $84,31 \pm 8,70\%$ ), os resultados obtidos foram comparáveis aos reportados por Minhas et al. (2021) (ACC: 84,9), Syaifullah et al. (2021) (ACC: 85,0) e Rahim et al. (2025) (ACC: 91,77). Além disso, superaram o desempenho apresentado por Xiao et al. (2021) (ACC: 75,9), bem como os demais trabalhos correlatos baseados em aprendizagem profunda (GAO et al., 2020; BAE et al., 2021; KUNG et al., 2021; ASHTARI-MAJLAN; SEIFI; DEHSHIBI, 2022; LUO et al., 2023).

Quanto à especificidade do Experimento 19 (Espec.:  $84,62 \pm 13,01\%$ ), os resultados foram inferiores aos obtidos por Minhas et al. (2021) (Espec.: 97,1%) e Syaifullah et al. (2021) (Espec.: 87,5%), assim como aos trabalhos baseados em aprendizagem profunda de Kung et al. (2021) (Espec.: 85,46%), Ashtari-Majlan, Seifi e Dehshibi (2022) (Espec.: 99,70%) e Rahim et al. (2025) (Espec.: 88,53%). No entanto, superaram os resultados de cinco outros estudos da literatura (GAO et al., 2020; SHU et al., 2021; XIAO et al., 2021; CALANDRELLI et al., 2022; LUO et al., 2023).

Quanto à sensibilidade do Experimento 19 (Sens.:  $84,00 \pm 9,67\%$ ), o resultado foi inferior apenas ao reportado por Rahim et al. (2025) (Sens.: 87,69%), sendo próximo ao de Syaifullah et al. (2021) (Sens.: 83,8%) e superior ao de nove outros trabalhos da literatura (GAO et al., 2020; KUNG et al., 2021; SHU et al., 2021; XIAO et al., 2021; MINHAS et al., 2021; ASHTARI-MAJLAN; SEIFI; DEHSHIBI, 2022; CALANDRELLI et al., 2022;

LUO et al., 2023).

Comparativamente, alguns modelos da literatura apresentaram desempenhos superiores ao obtido neste trabalho. No entanto, diversos fatores metodológicos podem ter influenciado esses resultados. O estudo de Minhas et al. (2021), por exemplo, alcançou métricas elevadas ao utilizar uma combinação de atributos, incluindo volume, área de superfície, espessura cortical e medidas neuropsicológicas em uma análise longitudinal. No entanto, a amostra utilizada foi bastante limitada — apenas 70 imagens de CCLe e 49 de CCLp — o que pode ter favorecido o *overfitting* e inflacionado os indicadores de desempenho. Por sua vez, Syaifullah et al. (2021) utilizou um conjunto de dados maior (128 imagens de CCLe e 284 de CCLp), mas restringiu-se a medidas de volume e avaliações neuropsicológicas, o que pode limitar a compreensão mais aprofundada da progressão da DA. Já o trabalho de Xie et al. (2023), embora tenha reportado uma AUC mais elevada e tenha utilizado medidas de volume, espessura e biomarcadores plasmáticos, não apresentou métricas de sensibilidade e especificidade, que são essenciais no contexto de diagnóstico médico, inviabilizando uma comparação completa com os demais métodos.

No caso de Rahim et al. (2025), apesar do bom desempenho reportado, o estudo utilizou menos de 100 imagens por grupo diagnóstico. Além disso, a aplicação de arquiteturas de redes neurais convolucionais profundas compromete a interpretabilidade do modelo, permitindo identificar as regiões envolvidas, mas não as características específicas que diferenciam CCLe de CCLp.

Em contrapartida, os modelos propostos neste trabalho foram treinados com uma base de dados mais robusta, composta por 264 imagens de CCLe e 247 de CCLp, o que garante maior representatividade dos grupos e contribui para uma avaliação mais realista e confiável do desempenho dos classificadores. Ademais, o modelo incorporou três medidas estruturais distintas — volume, espessura cortical e textura —, ampliando a capacidade de caracterização morfológica cerebral e contribuindo para uma compreensão mais abrangente da progressão da DA. Dessa forma, os modelos aqui desenvolvidos alcançam um equilíbrio importante entre desempenho preditivo e interpretabilidade, aspecto fundamental para aplicações clínicas.

## 5.6 Considerações finais

A acurácia e o equilíbrio entre classes alcançados pelo modelo proposto posicionam-no entre os classificadores mais robustos da literatura, com valores de sensibilidade e especificidade que asseguram uma detecção confiável tanto de casos de CCLe quanto de CCLp. Tal equilíbrio é particularmente relevante para aplicações clínicas, onde é essencial minimizar tanto os falsos positivos quanto os falsos negativos (RAHIM et al., 2025).

Além de seu desempenho competitivo, este estudo apresenta contribuições metodológicas significativas em relação a abordagens previamente reportadas. A utilização de uma

base de dados mais extensa e balanceada — composta por 264 imagens de CCLe e 247 de CCLp — amplia a representatividade amostral e confere maior realismo à avaliação do desempenho dos classificadores, reduzindo o risco de *overfitting*, comumente observado em estudos com amostras restritas. Ademais, a incorporação de três medidas morfológicas complementares (volume, espessura cortical e textura) permite uma caracterização estrutural cerebral mais abrangente, superando abordagens mais limitadas que se restringem a métricas volumétricas ou dados clínico-cognitivos.

Destaca-se, ainda, a adoção de técnicas de harmonização dos atributos estruturais, que visam mitigar variações sistemáticas introduzidas por diferentes *scanners* e protocolos de aquisição, reduzindo vieses indesejados e aumentando a comparabilidade entre os pacientes da base de dados.

Outro diferencial importante reside na interpretabilidade dos modelos gerados. A presente abordagem foi desenvolvida e validada exclusivamente com dados de uma única modalidade (ressonância magnética estrutural), utilizando atributos radiômicos e morfométricos derivados diretamente das imagens de RM. Além disso, a combinação de estratégias de seleção de atributos com classificadores de fácil compreensão clínica possibilita não apenas a identificação das regiões cerebrais mais discriminativas entre CCLe e CCLp, mas também das propriedades morfológicas específicas subjacentes a essa distinção. Esta abordagem representa um avanço em relação a modelos baseados em redes neurais profundas, os quais, embora frequentemente superiores em desempenho, carecem de transparência quanto aos mecanismos fisiopatológicos envolvidos, dificultando sua aplicação prática em cenários clínicos.

A avaliação rigorosa dos modelos propostos também merece destaque. Foram utilizadas estratégias robustas de validação cruzada e reportadas métricas complementares — incluindo AUC, acurácia, sensibilidade e especificidade —, o que fortalece a confiabilidade e a generalização dos resultados. Em conjunto, tais aspectos conferem ao modelo desenvolvido um equilíbrio fundamental entre desempenho preditivo e interpretabilidade, atributo indispensável para sua adoção em ambientes clínicos.

É importante ressaltar que diversos estudos com desempenho elevado reportado na literatura utilizaram conjuntos de dados consideravelmente menores — muitas vezes com menos de 100 amostras por classe —, o que suscita preocupações quanto ao *overfitting* e à limitação da capacidade de generalização (SYAIFULLAH et al., 2021; ASHTARI-MAJLAN; SEIFI; DEHSHIBI, 2022). Outros estudos empregaram amplificação extensiva de dados (*data augmentation*) ou integraram informações de modalidades que nem sempre estão disponíveis na prática clínica (MINHAS et al., 2021; RAHIM et al., 2025).

Adicionalmente, observa-se que o AUC elevado reportado em (ASHTARI-MAJLAN; SEIFI; DEHSHIBI, 2022) pode ter sido influenciado pelo uso intensivo de técnicas de amplificação de dados. Embora tais estratégias sejam eficazes em contextos de dados escassos, elas podem comprometer a comparabilidade com abordagens que empregam

atributos morfológicos extraídos manualmente e análises em volume completo, como a aqui apresentada.

Por fim, foi identificada uma limitação relevante relacionada à segmentação das regiões dos giros cerebrais, possivelmente decorrente da complexidade morfológica local e da baixa nitidez das fronteiras entre tecidos. Tais imprecisões podem impactar negativamente a acurácia na extração dos atributos estruturais e, por conseguinte, o desempenho dos modelos preditivos. Investimentos futuros no aprimoramento dessas segmentações — por meio de algoritmos mais sofisticados ou atlas anatômicos de alta resolução — têm o potencial de elevar significativamente a qualidade dos dados e, conseqüentemente, a sensibilidade e especificidade dos modelos para a detecção da progressão da doença.

O próximo capítulo apresentará as principais conclusões decorrentes desta investigação.



---

# Capítulo 6

## Conclusões

---

*Este capítulo sintetiza as principais conclusões do estudo do estado da arte em técnicas para a classificação de CCLp e CCLe.*

### 6.1 Conclusões

Com base nas análises conduzidas, este trabalho apresentou uma estrutura abrangente e metodologicamente robusta para a predição da conversão de CCL para DA, com ênfase na interpretabilidade e aplicabilidade clínica. O modelo desenvolvido, baseado em Floresta Aleatória, incorporou atributos radiômicos derivados de imagens de ressonância magnética estrutural, abrangendo medidas volumétricas, de espessura cortical e de textura, extraídas de diferentes regiões cerebrais. Essa diversidade de métricas permitiu capturar alterações morfológicas sutis associadas à progressão da DA, especialmente entre os CCLe e CCLp.

A utilização de um conjunto de dados relativamente grande e equilibrado (264 casos de CCLe e 247 de CCLp) contribuiu para aumentar a representatividade amostral e reduzir o risco de *overfitting*, frequentemente observado em estudos com amostras pequenas. Além disso, o uso de técnicas de harmonização, aplicadas aos atributos estruturais, foi eficaz na mitigação de efeitos de lote, melhorando a comparabilidade entre imagens adquiridas em diferentes centros.

Outro diferencial relevante do presente estudo foi o foco na interpretabilidade do modelo. Por meio da seleção de atributos e da aplicação de métodos explicativos baseados em SHAP, foi possível não apenas identificar as regiões cerebrais mais relevantes para a predição, mas também compreender quais atributos estruturais contribuíram para o desempenho do modelo. Essa abordagem contrasta com métodos de aprendizado profundo, que apesar de altamente precisos, geralmente carecem de transparência e dificultam a compreensão dos mecanismos envolvidos na progressão da doença.

Do ponto de vista metodológico, a adoção de validação cruzada aninhada e a utilização de métricas abrangentes de avaliação (acurácia, AUC, sensibilidade, especificidade e F1-score) garantiram uma estimativa confiável e robusta do desempenho do modelo. A combinação entre desempenho competitivo e capacidade de interpretação representa um passo importante em direção ao desenvolvimento de ferramentas clínicas de apoio à decisão.

Além disso, a relevância desta pesquisa alinha-se ao contexto das recentes políticas públicas brasileiras voltadas ao enfrentamento das demências, em especial à Lei nº 14.878/2024 (Brasil, 2024), que institui a “Política Nacional de Enfrentamento da Doença de Alzheimer e outras demências”. Entre seus princípios, destacam-se o estímulo à pesquisa científica e o incentivo ao desenvolvimento de ferramentas que viabilizem o diagnóstico precoce. Nesse sentido, o modelo proposto neste estudo representa um avanço concreto nessa direção. Ao oferecer uma abordagem acessível, interpretável e reproduzível, o método pode apoiar diretamente os esforços clínicos e institucionais voltados à identificação precoce de indivíduos com maior risco de conversão para a DA, contribuindo para intervenções mais eficazes.

Contudo, algumas limitações ainda persistem. A segmentação de regiões girais, devido à sua complexidade anatômica e limites pouco definidos, pode impactar a extração precisa dos atributos. A ausência de uma avaliação quantitativa da acurácia dessas segmentações impede uma análise mais rigorosa do impacto desse fator sobre os resultados obtidos. Futuras pesquisas deverão empregar técnicas de segmentação mais refinadas, como atlas de alta resolução ou métodos baseados em aprendizado profundo, visando melhorar a acurácia em regiões cerebrais desafiadoras.

Além disso, como os dados utilizados neste estudo foram obtidos exclusivamente a partir da base ADNI, composta por imagens adquiridas em centros norte-americanos, é importante ressaltar a necessidade de validar o método proposto em outras bases de dados, preferencialmente provenientes de diferentes populações e centros de aquisição, a fim de avaliar sua generalização e robustez em contextos clínicos mais diversos.

Em conclusão, este estudo contribui para o avanço do estado da arte na predição da conversão de CCL para DA, oferecendo uma abordagem interpretável, reproduzível e baseada em dados estruturais. Ele estabelece uma base sólida para futuras investigações que poderão incorporar dados longitudinais, multimodais e genéticos, ampliando a aplicabilidade e precisão dos modelos preditivos. A validação externa em outras bases de dados e o acompanhamento a longo prazo dos pacientes também serão etapas essenciais para consolidar o potencial clínico desta abordagem.

---

# Capítulo 7

## Apêndice

---

### 7.1 Apêndice 1. Atributos de textura de GLCM

As características de textura extraídas da matriz de coocorrência de níveis de cinza (GLCM) baseiam-se em medidas estatísticas de segunda ordem que descrevem as relações espaciais entre as intensidades dos voxels na imagem. Abaixo está um resumo de características selecionadas e suas interpretações:

**Cluster Shade** quantifica a assimetria da distribuição dos pares de níveis de cinza na textura. Valores mais altos indicam padrões de textura mais assimétricos e irregulares. É definido conforme a Equação 23:

$$\sum_{i,j} [(i - \mu_x) + (j - \mu_y)]^3 \cdot g(i, j), \quad (23)$$

sendo  $\mu_x$  e  $\mu_y$  são as médias das distribuições marginais.

**Probabilidade Máxima** captura a maior frequência de um par específico de níveis de cinza. Uma alta máxima probabilidade sugere uma textura repetitiva ou altamente estruturada. É expressa conforme a Equação 24:

$$\max (g(i, j)). \quad (24)$$

**Média da Soma** mede a média da soma dos níveis de cinza dos pares de pixels. Valores maiores podem refletir a dominância de pares de voxels de alta intensidade na textura. É definida conforme a Equação 25:

$$\frac{\sum_{i,j} (i + j) \cdot g(i, j)}{\sum_{i,j} g(i, j)}. \quad (25)$$

**Correlação da Informação Mútua I** (*Mutual Information Correlation I* - MIC I) mede a quantidade de informação compartilhada e é maior quando as dependências entre os níveis de cinza são mais fortes. É calculada conforme a Equação 26:

$$\frac{H(i, j) - H(i) - H(j)}{\max(H(i), H(j))}, \quad (26)$$

em que  $H(i, j)$  é a entropia conjunta e  $H(i)$ ,  $H(j)$  são as entropias marginais.

**MIC II**, uma medida complementar à MIC I, avalia a independência condicional por meio de uma transformação não linear. Valores mais altos indicam texturas mais simples, com menor interdependência entre os níveis de cinza. É expressa conforme Equação 27:

$$\sqrt{1 - \frac{e^{-\text{MIC}_I} - 1}{e^{\text{MIC}_I} - 1}}. \quad (27)$$

**Energia** captura a uniformidade local da textura. Maior energia implica em texturas mais homogêneas, com baixa variação nos níveis de cinza. É definida conforme a Equação 28:

$$\sum_{i,j} g(i, j)^2. \quad (28)$$

**Entropia** quantifica a aleatoriedade na distribuição dos níveis de cinza. Maior entropia está associada a texturas mais complexas ou desordenadas. É calculada conforme a Equação 29:

$$\sum_{i,j} g(i, j) \log_2 g(i, j). \quad (29)$$

**Entropia da Diferença** avalia a imprevisibilidade nas diferenças absolutas dos níveis de cinza. Valores altos refletem um contraste de textura mais diversificado. É expressa conforme a Equação 30:

$$-\sum_d p(d) \log(p(d)), \quad (30)$$

sendo  $p(d)$  é a probabilidade das diferenças nos níveis de cinza.

**Correlação** estima a dependência linear entre os níveis de cinza. É sensível a estruturas repetitivas alinhadas em direções específicas. É definida conforme a Equação 31:

$$\sum_{i,j} \frac{(i - \mu_x)(j - \mu_y)g(i, j)}{\sigma_x \cdot \sigma_y}. \quad (31)$$

**Média da Diferença** fornece uma medida do contraste médio entre voxels vizinhos. Valores maiores podem indicar maior variação local de intensidade. É calculada conforme a Equação 32:

$$\frac{\sum_{i,j} |i + j| \cdot g(i, j)}{\sum_{i,j} g(i, j)}. \quad (32)$$

**Diferença Inversa Normalizada** (*Inverse Difference Normalized* - IDN) avalia a homogeneidade da textura penalizando grandes diferenças de intensidade de forma normalizada. Valores mais altos indicam maior similaridade entre voxels vizinhos. É definida

conforme a Equação 33:

$$\sum_{i,j} \frac{g(i,j)}{1 + \left| \frac{i-j}{N_g} \right|}, \quad (33)$$

sendo  $N_g$  é o número de níveis de cinza.

**Momento da Diferença Inversa Normalizado** (*Inverse Difference Moment Normalized* - IDMN) aplica uma penalização mais forte para diferenças de intensidade, produzindo valores altos quando a textura é uniforme e as transições de intensidade são mínimas. É expressa conforme a Equação 34:

$$\sum_{i,j} \frac{g(i,j)}{1 + \left( \frac{i-j}{N_g} \right)^2}. \quad (34)$$

**Autocorrelação** quantifica a dependência linear dos níveis de cinza nos voxels vizinhos, servindo como indicador de padrões repetitivos ou regularidade na textura. Valores maiores sugerem maior grau de redundância estrutural. É definida conforme a Equação 35:

$$\sum_{i,j} i \cdot j \cdot p(i,j). \quad (35)$$

**Proeminência do Agrupamento** (*Cluster Prominence*) avalia o momento de quarta ordem da matriz de coocorrência dos níveis de cinza, capturando o grau de assimetria e a acuidade da distribuição dos níveis de cinza. Valores elevados estão associados a padrões texturais altamente enviesados e pontiagudos. É definida conforme a Equação 36:

$$\sum_{i,j} [(i - \mu_x) + (j - \mu_y)]^4 \cdot p(i,j). \quad (36)$$

**Tendência do Agrupamento** (*Cluster Tendency*) mede a tendência dos pixels com níveis de cinza semelhantes a se agruparem, oferecendo *insights* sobre a granulidade e a compactação da textura. Valores maiores indicam agrupamentos mais homogêneos. É definida conforme a Equação 37:

$$\sum_{i,j} [(i - \mu_x) + (j - \mu_y)]^2 \cdot p(i,j). \quad (37)$$

**Contraste** avalia a variação local de intensidade ao enfatizar as diferenças entre níveis de cinza vizinhos. É particularmente sensível à presença de bordas ou transições fortes de intensidade. É definida conforme a Equação 38:

$$\sum_{i,j} (i - j)^2 \cdot p(i,j). \quad (38)$$

**Variância da Diferença** quantifica a variância das diferenças entre pares de níveis de cinza, refletindo o grau de heterogeneidade na textura local. Valores maiores indicam aumento da variação local. É definida conforme a Equação 39:

$$\text{Var}(p_{|i-j|}). \quad (39)$$

**Diferença Inversa** (*Inverse Difference - ID*) caracteriza a homogeneidade da textura ao atribuir pesos maiores a pares de voxels com intensidades semelhantes. Fornece uma medida de uniformidade local. É definida conforme a Equação 40:

$$\sum_{i,j} \frac{p(i,j)}{1 + |i - j|}. \quad (40)$$

**Momento da Diferença Inversa** (*Inverse Difference Moment - IDM*) aumenta a sensibilidade a pequenas diferenças nos níveis de cinza aplicando uma penalidade quadrática às disparidades de intensidade dos voxels, sendo um indicador robusto de homogeneidade local. É definida conforme a Equação 41:

$$\sum_{i,j} \frac{p(i,j)}{1 + (i - j)^2}. \quad (41)$$

**Variância Inversa** é um refinamento do IDM, excluindo elementos diagonais e atribuindo penalidades maiores a diferenças maiores nos níveis de cinza. É definida conforme a Equação 42:

$$\sum_{i \neq j} \frac{p(i,j)}{(i - j)^2}. \quad (42)$$

**Média Conjunta** (*Joint Average*) representa a intensidade média dos pares de voxels coocorrentes, oferecendo um resumo global da distribuição dos níveis de cinza entre voxels espacialmente correlacionados. É definida conforme a Equação 43:

$$\sum_{i,j} \frac{(i + j)}{2} \cdot p(i,j). \quad (43)$$

**Coefficiente Máximo de Correlação** (*Maximal Correlation Coefficient - MCC*) estima a complexidade da textura ao quantificar a maior correlação possível entre linhas e colunas em uma matriz derivada da GLCM. Valores elevados de MCC refletem dependências estruturais complexas. É definida conforme a Equação 44:

$$\sqrt{\lambda_2}, \quad (44)$$

sendo  $\lambda_2$  o segundo maior autovalor de  $Q$ , conforme a Equação 45:

$$Q(i,j) = \sum_k \frac{p(i,k) \cdot p(j,k)}{p(i) \cdot p(j)}. \quad (45)$$

**Entropia da Soma** avalia a aleatoriedade associada à soma dos níveis de cinza em pares de voxels, servindo como uma medida da desordem global. Valores maiores indicam maior complexidade distributiva. É definida conforme a Equação 46:

$$- \sum_{k=2}^{2N_g} p_{x+y}(k) \cdot \log(p_{x+y}(k)). \quad (46)$$

**Quadrados da Soma** calcula a dispersão dos níveis de cinza em torno da média, oferecendo uma medida de contraste e variabilidade textural. É definida conforme a Equação 47:

$$\sum_{i,j} (i - \mu)^2 \cdot p(i,j). \quad (47)$$

## 7.2 Apêndice 2. Abreviações usadas nas figuras

A Tabela 20 apresenta as abreviações e as descrições correspondentes dos vinte principais atributos de imagem ilustrados nos gráficos das Figuras 48 e 49, a fim de facilitar a compreensão.

Tabela 20 – Abreviações de atributos e descrições correspondentes.

<b>Abreviação</b>	<b>Descrição</b>
Vol-GM-D-Amig	Volume de GM da amígdala direita
ECMéd-E-Amig	Espessura cortical média da amígdala esquerda
ECMéd-D-Amig	Espessura cortical média da amígdala direita
Vol-GM-D-Gir-Temp-Inf	Volume de GM do giro temporal inferior direito
Vol-WM-E-Amig	Volume de WM da amígdala esquerda
Vol-WM-E-Nuc-Accum	Volume de WM do núcleo accumbens esquerdo
Vol-GM-E-Amig	Volume de GM da amígdala esquerda
Vol-WM-D-Amig	Volume de WM da amígdala direita
IDMN-D-Hipo	Atributo de textura IDMN do hipocampo direito
Vol-GM-D-Gir-Temp-Méd	Volume de GM do giro temporal médio direito
IDN-D-Amig	Atributo de textura IDN da amígdala direita
Vol-WM-D-Gir-Fus	Volume de WM do giro fusiforme direito
Vol-GM-E-Gir-Temp-Méd	Volume de GM do giro temporal médio esquerdo
SomMéd-E-Nuc-Accum	Atributo de textura Soma Média do núcleo accumbens esquerdo
IDMN-D-Amig	Atributo de textura IDMN da amígdala direita
Vol-D-Gir-Temp-Inf	Volume do giro temporal inferior direito
ECMéd-D-Polo-Temp	Espessura cortical média do polo temporal direito
Energia-D-Amig	Atributo de textura Energia da amígdala direita
Vol-WM-E-Polo-Temp	Volume de WM do polo temporal esquerdo
DifMéd-E-Gir-Temp-Sup	Atributo de textura Diferença Média do giro temporal superior esquerdo

Fonte: Autoria própria.

---

## Referências

---

ADNI. **About ADNI**. 2024. Acessado em: 11 de abril de 2024. Disponível em: <<https://adni.loni.usc.edu/about/>>.

AJA-FERNÁNDEZ, S.; VEGAS-SÁNCHEZ-FERRERO, G. The Problem of Noise in MRI. In: **Statistical Analysis of Noise in MRI: Modeling, Filtering and Estimation**. Switzerland: Springer Nature, 2016. cap. 1, p. 1–6.

ALI, J. et al. Random forests and decision trees. **International Journal of Computer Science Issues (IJCSI)**, v. 9, n. 5, p. 272–278, 2012.

Alzheimer's Association. 2023 Alzheimer's disease facts and figures. **Alzheimer's & Dementia**, v. 19, p. 1598–1695, 2023.

\_\_\_\_\_. **Alzheimer's Disease Graphic**. 2024. Acessado em: 13 de fevereiro de 2024. Disponível em: <<https://medicine.umich.edu/dept/mneuronet/alzheimers-disease-graphic>>.

Alzheimer's Disease International. **World Alzheimer Report 2023**. 2023. <https://www.alzint.org/resource/world-alzheimer-report-2023/>. Acessado em: 27 de novembro de 2023.

AMARAL, D. G.; STRICK, P. L. A organização do sistema nervoso central. In: KANDEL, E. R. et al. (Ed.). **Princípios de neurociências**. 5. ed. Porto Alegre: AMGH Editora Ltda., 2014. cap. 15, p. 299 – 314.

ASHTARI-MAJLAN, M.; SEIFI, A.; DEHSHIBI, M. M. A multi-stream convolutional neural network for classification of progressive MCI in Alzheimer's disease using structural MRI images. **IEEE Journal of Biomedical and Health Informatics**, IEEE, v. 26, n. 8, p. 3918–3926, 2022.

ATRI, A. The Alzheimer's disease clinical spectrum: diagnosis and management. **Medical Clinics**, Elsevier, v. 103, n. 2, p. 263–293, 2019.

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.

AZAMBUJA, L. Avaliação neuropsicológica do idoso. **Revista Brasileira De Ciências Do Envelhecimento Humano**, v. 4, n. 2, p. 40–45, 2007.

- BAE, J. et al. Transfer learning for predicting conversion from mild cognitive impairment to dementia of Alzheimer's type based on a three-dimensional convolutional neural network. **Neurobiology of aging**, Elsevier, v. 99, p. 53–64, 2021.
- BESAG, J. On the statistical analysis of dirty pictures. **Journal of the Royal Statistical Society Series B (Methodological)**, v. 48, n. 3, p. 259–279, 1986.
- Brasil. **Lei nº 14.878, de 4 de junho de 2024**. 2024. Diário Oficial da União, Edição 106, Seção 1, p. 1. Disponível em: <[https://www.planalto.gov.br/ccivil\\_03/\\_Ato2023-2026/2024/Lei/L14878.htm](https://www.planalto.gov.br/ccivil_03/_Ato2023-2026/2024/Lei/L14878.htm)>.
- BRUCKI, S. M. D.; CARAMELLI, P.; NITRINI, R. Demências. In: NITRINI, R.; BACHESCHI, L. A. (Ed.). **A Neurologia que todo médico deve saber**. 3. ed. São Paulo: Editora Atheneu, 2015. cap. 15, p. 337–354.
- BRUNDEN, K. R.; TROJANOWSKI, J. Q.; LEE, V. M.-Y. Advances in Tau-focused drug discovery for Alzheimer's disease and related tauopathies. **Nature Reviews Drug Discovery**, v. 8, n. 10, p. 783–793, 2009.
- BUADES, A.; COLL, B.; MOREL, J.-M. A non-local algorithm for image denoising. In: IEEE. **2005 IEEE computer society conference on computer vision and pattern recognition**. [S.l.], 2005. v. 2, p. 60–65.
- \_\_\_\_\_. A review of image denoising algorithms, with a new one. **Multiscale modeling & simulation**, SIAM, v. 4, n. 2, p. 490–530, 2005.
- BUSHBERG, J. T. et al. **The Essential Physics of Medical Imaging**. 3. ed. Philadelphia, USA: Lippincott Williams & Wilkins, 2012.
- CAGNIE, B. et al. Muscle Functional MRI as an Imaging Tool to Evaluate Muscle Activity. **Journal of Orthopaedic & Sports Physical Therapy**, v. 41, n. 11, p. 896–903, 2011.
- CALANDRELLI, R. et al. Brain atrophy pattern in patients with mild cognitive impairment: MRI study. **Translational neuroscience**, De Gruyter, v. 13, p. 335–348, 2022.
- da COSTA, M. C. et al. **Histologia Interativa: Tecido Nervoso**. 2019. Disponível em: <<https://www.unifal-mg.edu.br/histologiainterativa/tecido-nervoso/>>.
- de ALVARENGA JÚNIOR, W. J. **Métodos de otimização hiperparamétrica: um estudo comparativo utilizando árvores de decisão e florestas aleatórias na classificação binária**. Dissertação (Mestrado) — Universidade Federal de Minas Gerais, Belo Horizonte, 2018.
- de MENDONÇA, L. J. C.; FERRARI, R. J. Alzheimer's disease classification based on graph kernel SVMs constructed with 3D texture features extracted from MR images. **Expert Systems with Applications**, v. 211, p. 118633, 2023. ISSN 0957-4174.
- DINIZ, G. L.; UTINO, B. M. Doença de Alzheimer: Uma revisão da epidemiologia, diagnóstico e tratamento. In: NETO, B. R. D. S. (Ed.). **Neurologia: investigação clínica e o estudo dos mecanismos das doenças**. Ponta Grossa: Editora Atena, 2023. cap. 3, p. 18–27.

- ELIAS JUNIOR, J. et al. Complicações do uso intravenoso de agentes de contraste à base de gadolínio para ressonância magnética. **Radiologia Brasileira**, v. 41, n. 4, p. 263–267, 2008.
- FACELI, K. et al. Métodos Baseados em Procura. In: **Inteligência artificial: uma abordagem de aprendizado de máquina**. Rio de Janeiro: LTC, 2011. p. 83–106.
- FOSS, M. P.; VALE, F. d. A. C. d.; SPECIALI, J. G. Influência da escolaridade na avaliação neuropsicológica de idosos: Aplicação e análise dos resultados da Escala de Mattis para Avaliação de Demência ( Mattis Dementia Rating Scale - MDRS). **Arquivos de Neuro-Psiquiatria**, SciELO Brasil, v. 63, n. 1, p. 119–126, 2005.
- FREIRE, P. G. d. L. **Automatic computational scheme for segmentation, volumetric assessment and analysis of multiple sclerosis lesions in magnetic resonance images of the human brain**. Tese (Doutorado) — FEDERAL UNIVERSITY OF SÃO CARLOS, São Carlos - SP - Brasil, 2019.
- GAO, F. et al. AD-NET: Age-adjust neural network for improved MCI to AD conversion prediction. **NeuroImage: Clinical**, Elsevier, v. 27, p. 102290, 2020.
- GISPERT, J. D. et al. Method for bias field correction of brain T1-weighted Magnetic Resonance Images minimizing segmentation error. **Human Brain Mapping**, v. 22, n. 2, p. 133–144, 2004.
- GUAN, H. et al. Attention-guided autoencoder for automated progression prediction of subjective cognitive decline with structural MRI. **IEEE Journal of Biomedical and Health Informatics**, IEEE, 2023.
- HAGE, M. C. F. N. S.; IWASAKI, M. Imagem por ressonância magnética: princípios básicos. **Ciência Rural**, v. 39, n. 4, p. 1287–1295, 2009.
- HALLE, M. et al. **Slagter - Drawing Coronal section of the brain - no labels**. 2017. Acessado em: 6 de maio de 2024. Disponível em: <<https://www.openanatomy.org/atlas-pages/atlas-spl-brain.html>>.
- HANSON, L. G.; GROTH, T. Introduction to Magnetic Resonance Imaging Techniques. In: **Clinical and research applications of diagnostic imaging techniques: MR, PET, SPECT, CT and ultrasound: PhD Course**. Dinamarca: DTU Orbit, 2009. p. 48.
- IGLESIAS, J. E. et al. Robust brain extraction across datasets and comparison with publicly available methods. **IEEE Transactions on Medical Imaging**, IEEE, v. 30, n. 9, p. 1617–1634, 2011.
- ISSA, I. **The Use of Metasurfaces for Magnetic Resonance Imaging Applications**. Tese (Doutorado) — University of Sheffield, 2018.
- JACOBS, M. A.; IBRAHIM, T. S.; OUWERKERK, R. MR Imaging: Brief Overview and Emerging Applications. **RadioGraphics**, v. 27, n. 4, p. 1213–1229, 2007.
- JOHNSON, H. J.; MCCORMICK, M. M.; IBÁÑEZ, L. **The ITK Software Guide**. 4. ed. [S.l.], 2021. Accessed: 2024-05-20. Disponível em: <<https://docs.itk.org/en/latest/>>.

- JOHNSON, W. E.; LI, C.; RABINOVIC, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. **Biostatistics**, Oxford University Press, v. 8, n. 1, p. 118–127, 2007.
- KUNG, T.-H. et al. Neuroimage biomarker identification of the conversion of mild cognitive impairment to Alzheimer’s Disease. **Frontiers in Neuroscience**, Frontiers Media SA, v. 15, n. 584641, 2021.
- LEANDROU, S. et al. Comparison of 1.5 T and 3 T MRI hippocampus texture features in the assessment of Alzheimer’s disease. **Biomedical Signal Processing and Control**, Elsevier, v. 62, p. 102098, 2020.
- LEE, S.; KIM, K. W.; for the Alzheimer’s Disease Neuroimaging Initiative. Associations between texture of T1-weighted magnetic resonance imaging and radiographic pathologies in Alzheimer’s disease. **European Journal of Neurology**, Wiley Online Library, v. 28, n. 3, p. 735–744, 2021.
- LEE, S.; LEE, H.; KIM, K. W. Magnetic resonance imaging texture predicts progression to dementia due to Alzheimer disease earlier than hippocampal volume. **Journal of Psychiatry and Neuroscience**, Journal of Psychiatry and Neuroscience, v. 45, n. 1, p. 7–14, 2020.
- LI, C.; GORE, J. C.; DAVATZIKOS, C. Multiplicative intrinsic component optimization (MICO) for MRI bias field estimation and tissue segmentation. **Magnetic Resonance Imaging**, v. 32, n. 7, p. 913–923, 2014.
- LI, Y. et al. Impact of preprocessing and harmonization methods on the removal of scanner effects in brain MRI radiomic features. **Cancers**, MDPI, v. 13, n. 12, p. 3000, 2021.
- LIN, S.-Y. et al. The clinical course of early and late mild cognitive impairment. **Frontiers in Neurology**, Frontiers Media SA, v. 13, p. 685636, 2022.
- LÓPEZ, M. M. M.; FREDERICK, J. M.; VENTURA, J. Evaluation of MRI denoising methods using unsupervised learning. **Frontiers in Artificial Intelligence**, Frontiers Media SA, v. 4, n. 642731, 2021.
- LUNDBERG, S. M.; LEE, S.-I. A unified approach to interpreting model predictions. In: **31st Conference on Neural Information Processing Systems (NIPS 2017)**. Long Beach, CA, USA: Neural Information Processing Systems Foundation, 2017.
- LUO, M. et al. Class activation attention transfer neural networks for MCI conversion prediction. **Computers in Biology and Medicine**, Elsevier, v. 156, n. 106700, 2023.
- MALLIKA, R. et al. Texture Features Based Hybrid Multi Support Vector Machine Model for the diagnosis of Alzheimer’s disease through Brain MRI Images. In: **IEEE. 2023 International Conference on Computational Intelligence, Networks and Security (ICCINS)**. [S.l.], 2023. p. 1–6.
- MANJÓN, J. V. et al. Adaptive non-local means denoising of mr images with spatially varying noise levels. **Journal of Magnetic Resonance Imaging**, Wiley Online Library, v. 31, n. 1, p. 192–203, 2010.

- MARTELETTO, S. R. **Técnicas de seleção de atributos através de Random Forests: um estudo de caso para detecção de tendências em séries temporais financeiras**. Dissertação (Mestrado) — Universidade de São Paulo, São Paulo, 2022.
- 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, 2009.
- MINHAS, S. et al. Early MCI-to-AD Conversion Prediction Using Future Value Forecasting of Multimodal Features. **Computational intelligence and neuroscience**, Wiley Online Library, v. 2021, 2021.
- Ministério da Saúde. **Linhas de Cuidado - Demência**. 2024. Acessado em: 14 de fevereiro de 2024. Disponível em: <<https://linhasdecuidado.saude.gov.br/portal/demencia/atencao-especializada/rastreamento/>>.
- 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.
- NZIMBI, R. S. **Slagter - Drawing Coronal section of the brain - no labels**. 2024. Acessado em: 9 de fevereiro de 2024. Disponível em: <<https://anatomytool.org/content/slagter-drawing-coronal-section-brain-no-labels>>.
- PEI, L. et al. A general skull stripping of multiparametric brain MRIs using 3D convolutional neural network. **Scientific Reports**, Nature Publishing Group UK London, v. 12, n. 10826, 2022.
- POIRIER, J.; GAUTHIER, S. **Doença de Alzheimer: o guia completo**. [S.l.]: MG Editores, 2016.
- POOLEY, R. A. Fundamental Physics of MR Imaging. **RadioGraphics**, v. 25, n. 4, p. 1087–1099, 2005.
- PORSTEINSSON, A. et al. Diagnosis of early Alzheimer’s disease: clinical practice in 2021. **The Journal of Prevention of Alzheimer’s Disease**, Springer, v. 8, p. 371–386, 2021.
- Queensland Brain Institute. **Lobes of the Brain**. 2024. Acessado em: 9 de fevereiro de 2024. Disponível em: <<https://qbi.uq.edu.au/brain/brain-anatomy/lobes-brain>>.
- RAHIM, N. et al. Early progression detection from mci to ad using multi-view mri for enhanced assisted living. **Image and Vision Computing**, Elsevier, v. 157, p. 105491, 2025.
- REHMAN, H. Z. U.; HWANG, H.; LEE, S. Conventional and deep learning methods for skull stripping in brain MRI. **Applied Sciences**, MDPI, v. 10, n. 5, p. 1773, 2020.
- ROOZPEYKAR, S. et al. Contrast-enhanced weighted-T1 and FLAIR sequences in MRI of meningeal lesions. **American Journal of Nuclear Medicine and Molecular Imaging**, v. 12, n. 2, p. 63–70, 2022.
- SAJJADI, S. A.; BROWN, J. Clinical assessment of patients with dementia. **Advances in clinical neurology and rehabilitation**, v. 15, n. 10, p. 10–13, 2015.

SANVITO, W. L.; ROCHA, A. J. **Síndromes Neurológicas**. 3. ed. São Paulo: Editora Atheneu, 2008. 14–20 p.

SHU, Z.-Y. et al. Prediction of the progression from mild cognitive impairment to Alzheimer’s disease using a radiomics-integrated model. **Therapeutic Advances in Neurological Disorders**, SAGE Publications Sage UK: London, England, v. 14, p. 1–13, 2021.

SILVA, J. et al. Structural MRI texture analysis for detecting Alzheimer’s disease. **Journal of Medical and Biological Engineering**, Springer, v. 43, n. 3, p. 227–238, 2023.

SILVEIRA, R. V. D. et al. Differences between Alzheimer’s disease and mild cognitive impairment using brain networks from magnetic resonance texture analysis. **Experimental brain research**, Springer, v. 242, p. 1947–1955, 2024.

SLED, J. G.; ZIJDENBOS, A. P.; EVANS, A. C. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. **IEEE Transactions on Medical Imaging**, IEEE, v. 17, n. 1, p. 87–97, 1998.

SMID, J. et al. Declínio cognitivo subjetivo, comprometimento cognitivo leve e demência - diagnóstico sindrômico: recomendações do Departamento Científico de Neurologia Cognitiva e do Envelhecimento da Academia Brasileira de Neurologia. **Dementia & Neuropsychologia**, SciELO Brasil, v. 16, p. 1–17, 2022.

SYAIFULLAH, A. H. et al. Machine learning for diagnosis of AD and prediction of MCI progression from brain MRI using brain anatomical analysis using diffeomorphic deformation. **Frontiers in Neurology**, Frontiers Media SA, v. 11, p. 576029, 2021.

TIMMEREN, J. E. V. et al. Radiomics in medical imaging—“how-to” guide and critical reflection. **Insights into imaging**, Springer, v. 11, n. 1, p. 91, 2020.

TOFTS, P. **Quantitative MRI of the Brain: Measuring Changes Caused by Disease**. Inglaterra: John Wiley & Sons Ltd, 2003.

TORRES, K. C. d. L. et al. Biomarcadores na doença de Alzheimer. **Geriatrics, Gerontology and Aging**, v. 6, n. 3, p. 273–282, 2012.

TUSTISON, N. J. et al. N4ITK: Improved N3 Bias Correction. **IEEE Transactions on Medical Imaging**, IEEE, v. 29, n. 6, p. 1310–1320, 2010.

VELAZQUEZ, M.; LEE, Y. Random forest model for feature-based Alzheimer’s disease conversion prediction from early mild cognitive impairment subjects. **Plos one**, Public Library of Science San Francisco, CA USA, v. 16, n. 4, 2021.

WEST, H. et al. Clinical validation of synthetic brain MRI in children: initial experience. **Neuroradiology**, v. 59, n. 1, 2017.

XIAO, R. et al. Early diagnosis model of Alzheimer’s disease based on sparse logistic regression with the generalized elastic net. **Biomedical Signal Processing and Control**, Elsevier, v. 66, n. 102362, 2021.

XIE, L. et al. Baseline structural MRI and plasma biomarkers predict longitudinal structural atrophy and cognitive decline in early Alzheimer's disease. **Alzheimer's Research & Therapy**, Springer, v. 15, n. 79, 2023.

\_\_\_\_\_. Longitudinal atrophy in early Braak regions in preclinical Alzheimer's disease. **Human brain mapping**, Wiley Online Library, v. 41, n. 16, p. 4704–4717, 2020.

ZAIN, M. F. M.; MAHMUD, W. M. H. W.; GAN, H.-S. Preliminary Analysis on the Effect of Different Denoising Techniques towards Texture Features of MRI Images of Alzheimer's Disease. **Journal of Advanced Research in Applied Sciences and Engineering Technology**, v. 31, n. 2, p. 234–244, 2023.

ZHANG, J. et al. 3D texture analysis on MRI images of Alzheimer's disease. **Brain imaging and behavior**, Springer, v. 6, p. 61–69, 2012.



# 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 Vivian Ferreira Margarido, realizada em 24/07/2025.

### Comissão Julgadora:

Prof. Dr. Ricardo José Ferrari (UFSCar)

Prof. Dr. Alan Demétrius Baria Valejo (UFSCar)

Prof. Dr. Marcelo Zancheta do Nascimento (UFU)

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