

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

Lucas José Cruz de Mendonça

**Classificação da doença de Alzheimer
baseada em *graph kernels* construídos a
partir de atributos de textura 3D de
imagens de ressonância magnética**

São Carlos
Janeiro de 2022

Lucas José Cruz de Mendonça

**Classificação da doença de Alzheimer
baseada em *graph kernels* construídos a
partir de atributos de textura 3D de
imagens de ressonância magnética**

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

Área de concentração: Metodologias e Técnicas de Computação

Orientador: Prof. Dr. Ricardo José Ferrari

São Carlos
Janeiro de 2022



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 Lucas José Cruz de Mendonça, realizada em 16/02/2022.

Comissão Julgadora:

Prof. Dr. Ricardo José Ferrari (UFSCar)

Prof. Dr. Alexandre Luis Magalhães Levada (UFSCar)

Prof. Dr. Leandro Alves Neves (UNESP)

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

Agradecimentos

Agradeço ao meu orientador, Prof. Dr. Ricardo José Ferrari, por sua orientação, por compartilhar seus conhecimentos, pela presença constante e pelas discussões de ideias sempre produtivas.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES - pelo apoio financeiro (Código de Financiamento 001).

E aos meus pais, por todo o suporte e apoio.

Resumo

A doença de Alzheimer (DA) é neurodegenerativa e caracterizada pelo comprometimento cognitivo. O comprometimento cognitivo leve (CCL) é uma condição clínica que envolve um leve déficit de memória e em muitos casos representa um estado de transição entre uma condição cognitivamente normal (CN) e a DA. Estudos publicados na literatura restringem suas análises à busca por alterações estruturais causadas pela doença em algumas regiões do cérebro. Atualmente, as pesquisas estão voltadas à procura de novos biomarcadores da AD usando múltiplas regiões do cérebro e com foco em mudanças sutis de textura na imagem. Portanto, este estudo propõe uma nova técnica para a classificação de imagens de ressonância magnética estrutural (RM) no diagnóstico de DA usando *graph kernel* (GK). Neste método, primeiro segmentamos as imagens cerebrais em várias regiões com o FreeSurfer. Em seguida, extraímos 22 atributos de textura e definimos os atributos dos vértices dos grafos como as distribuições de probabilidade destes atributos. Para cada atributo de textura, construímos os grafos e definimos os pesos das arestas pela distância entre pares de atributos de vértices usando três métricas. A remoção de arestas dos grafos foi baseada em limiar. Por fim, realizamos a classificação dos grafos usando *Support Vector Machines* (SVMs) com dois GKs. Os resultados do nosso método mostraram melhores desempenhos para as classificações CN×AD (AUC=0.92) e CN×CCL (AUC=0.81), e piores para CCL×AD (AUC=0.78). Essa tendência é consistente com outros resultados publicados e faz sentido se considerarmos o conceito de continuum da DA do ponto de vista fisiopatológico e clínico. Além de permitir o uso de diferentes atributos de textura para o diagnóstico da DA, nosso método utiliza a abordagem GK para representar atributos de textura de diferentes regiões cerebrais o que facilita a tarefa de classificação de imagens via SVMs. Nossos resultados foram promissores quando comparados ao estado da arte na classificação de AD baseada em grafos.

Palavras-chave: Doença de Alzheimer. Classificação de Grafos. Extração de Textura. Graph Kernel. Imagem de Ressonância Magnética.

Abstract

Alzheimer’s disease (AD) is neurodegenerative and characterized by cognitive and behavioral impairment. Mild cognitive impairment (MCI) is a relatively broad clinical condition involving a slight memory deficit, which in many cases represents a transitional state between a cognitively normal (CN) condition and AD. Many published studies restrict their analyses to searching for structural changes caused by the disease in a few particular regions of the brain.. Currently, the studies are looking for new AD biomarkers using multiple brain regions and focusing on subtle texture changes in the image. Therefore, this study proposes a new technique for MR image classification in AD diagnosis using graph kernels constructed from texture features extracted from 3D Structural magnetic resonance (sMR) images. In our proposed method we first segment the brain images into multiple regions with the FreeSurfer. Then, we extract 22 texture features and defined the graph node attributes as the probability distributions of the extracted features. Next, for each texture feature we build a graph and define its edge weights as the distances between pairs of node attributes using three metrics. After that, we use a threshold-based approach for graph edges removal. Finally, we perform graph classification using Support Vector Machines (SVMs) with two graph-kernels. Results of our method have shown better performances for the CN×AD (AUC=0.92) and CN×MCI (AUC=0.81) classifications, and worse for MCI×AD (AUC=0.78). This trend is consistent with other published results and makes sense if we consider the concept of Alzheimer’s disease continuum from pathophysiological, biomarker and clinical perspectives. Besides allowing the use of different texture attributes for the diagnosis of AD, our method uses the graph-kernel approach to represent texture features from different regions of the brain image, which considerably facilitates the image classification task via SVMs. Our results were promising when compared to the state-of-the-art in graph-based AD classification.

Keywords: Alzheimer’s Disease. Graphs-based Classification. Texture Extraction. Graph Kernel. Magnetic Resonance Image.

Lista de ilustrações

Figura 1 – Diferenças estruturais entre um cérebro saudável (cognitivamente normal) e com doença de Alzheimer. Extraído e adaptado de (ALZHEIMER'S ASSOCIATION, 2020).	20
Figura 2 – Diagrama de blocos indicando o fluxo de funcionamento do projeto. . .	36
Figura 3 – Segmentação de sub regiões do hipocampo pelo algoritmo proposto por Iglesias et al. (2015).	38
Figura 4 – A representação visual das primeiras etapas da Figura 2 para um paciente.	45
Figura 5 – Exemplo de grafo G com atributos nos vértices.	49
Figura 6 – Exemplo de funcionamento do <i>kernel</i> P2K para dois grafos com atributos de uma dimensão.	54
Figura 7 – Esquema de classificação utilizado neste estudo.	56
Figura 8 – Ilustração de duas máscaras binárias dos hipocampus esquerdo e direito extraídas e alinhadas à imagem de original do paciente, em cortes (a) Axial (b) Sagital e (c) Coronal.	59
Figura 9 – Cortes sagitais de uma imagem de um dado paciente após o seu registro com o espaço do FreeSurfer.	60
Figura 10 – Ilustração das máscaras binárias obtidas pela segmentação da imagem do paciente apresentado na Figura 9 em um corte sagital.	61
Figura 11 – Dois atributos do GLCM em cortes sagitais de uma imagem do paciente apresentado na Figura 9.	61
Figura 12 – Dois atributos do RLM em cortes sagitais de uma imagem do paciente apresentado na Figura 9.	62
Figura 13 – Dois atributos do LBP em cortes sagitais de uma imagem do paciente apresentado na Figura 9	62

Figura 14 – Distribuições de textura (a) não normalizadas e (b) normalizadas para a região da amígdala esquerda de um paciente para cada classe diagnóstica. As distâncias de Wasserstein entre as distribuições normalizadas são $D_{ws} = 0,370$, $D_{ws} = 0,221$ e $D_{ws} = 0,148$ para os casos CCL×DA, CN×CCL e CN×DA, respectivamente.	63
Figura 15 – Distribuições de textura para o atributo F_{14} do RLM, para dois pacientes DA (DA 1 e DA 2). A região mais escura na imagem representa a sobreposição das duas distribuições.	64
Figura 16 – Ilustração do alinhamento de prótons paralelamente e antiparalelamente à um campo magnético B_0 , extraído e adaptado de (BUSHBERG; BOONE, 2011).	86
Figura 17 – Representação do processo de relaxamento dos prótons após a remoção do segundo campo magnético, extraído de (MAZZOLA, 2009).	86
Figura 18 – Ilustração do funcionamento dos hiperplanos no SVM, com amostras em diferentes valores de ξ	87

Lista de tabelas

Tabela 1 – Resumo dos trabalhos recentes relacionados ao tema dessa pesquisa. . .	32
Tabela 2 – Conjunto de imagens da base de dados ADNI.	37
Tabela 3 – Regiões segmentadas pelo atlas padrão do FreeSurfer.	39
Tabela 4 – Regiões segmentadas pelo atlas DKT40 do FreeSurfer.	39
Tabela 5 – Resultados de classificação para CN×DA. A acurácia (ACC), sensibilidade (SENS) e especificidade (SPEC) são dadas em porcentagem. . .	65
Tabela 6 – Resultados de classificação para CN×CCL. A acurácia (ACC), sensibilidade (SENS) e especificidade (SPEC) são dadas em porcentagem. .	66
Tabela 7 – Resultados de classificação para CCL×DA. A acurácia (ACC), sensibilidade (SENS) e especificidade (SPEC) são dadas em porcentagem. .	66
Tabela 8 – Comparação de resultados com outros métodos. A acurácia (ACC) é dada em porcentagem.	67

Lista de siglas

AAL	Automated Anatomical Labeling
ADNI	Alzheimers Disease Neuroimaging Initiative
AUC	Area Under Curve
CCLe	Comprometimento cognitivo leve estável
CCLP	Comprometimento cognitivo leve precoce
CCLpr	Comprometimento cognitivo leve progressivo
CCLT	Comprometimento cognitivo leve tardio
CN	Cognitivamente normal
CSF	Cerebral spinal fluid
DA	Doença de Alzheimer
fMRI	Functional Magnetic Resonance Imaging
GK	Graph Kernel
GLCM	Gray Level Co-Occurrence Matrix
GLN	Gray level non uniformity
GM	Gray matter
HGK	Hash Graph Kernel
HGRE	High gray level run emphasis
IDM	Inverse difference moment

kPCA	Kernel Principle Components Analysis
LBP	Local Binary Patterns
LGRE	Low gray level run emphasis
LRE	Long run emphasis
LRHGE	Long run high gray level emphasis
LRLGE	Long run low gray level emphasis
LSH	Locality Sensitive Hashing
MST	Minimum Spanning Tree
P2K	Propagation Kernel
PGB	Patch-Based Grading
RF	Random Forest
RLM	Run Length Matrix
RLN	Run length non uniformity
RM	Ressonância Magnética
ROIs	Regions of Interest (Regiões de interesse)
SP	Shortest-Path
SRE	Short run emphasis
SRHGE	Short run high gray level emphasis
SRLGE	Short run lowgray level emphasis
SVM	Support Vector Machine
WL	Weisfeiler-Lehman
WM	White matter
wSP	weighted SP

Sumário

1	INTRODUÇÃO	19
1.1	Contexto e motivação	19
1.2	Objetivos	22
1.2.1	Objetivos específicos	22
1.3	Organização do trabalho	23
2	REVISÃO BIBLIOGRÁFICA	25
2.1	Classificação de grafos em ressonância magnética	25
2.1.1	Trabalhos correlatos com <i>Graph Kernel</i>	25
2.1.2	Trabalhos correlatos sem <i>Graph Kernel</i>	28
2.1.3	Discussão e considerações finais	32
3	METODOLOGIA	35
3.1	Visão geral	35
3.2	Base de dados	36
3.3	Preprocessamento e segmentação	37
3.3.1	FreeSurfer	37
3.3.2	Métodos utilizados	37
3.4	Filtros de textura	40
3.4.1	GLCM	40
3.4.2	RLM	42
3.4.3	LBP	44
3.5	Extração de características de textura	44
3.5.1	Normalização	45
3.5.2	Histogramas e distribuições de texturas	46
3.5.3	Volume da região e momentos estatísticos	48
3.6	Construção dos grafos cerebrais	48

3.6.1	Atributos dos vértices	49
3.6.2	Definição das arestas e seus pesos	50
3.6.3	Limiarização	51
3.7	Graph Kernel	52
3.7.1	Propagation Kernel	53
3.7.2	Hash Graph Kernel	54
3.8	Esquema de classificação	56
3.9	Métricas de avaliação dos resultados	57
3.10	Considerações gerais da metodologia	57
4	RESULTADOS E DISCUSSÕES	59
4.1	Pipeline do FreeSurfer e segmentação	59
4.2	Filtros de textura	60
4.3	Extração de características e normalizações	62
4.4	Classificação e discussão	64
4.4.1	Comparação com métodos da literatura	67
4.5	Considerações finais	69
5	CONCLUSÕES	71
5.1	Conclusões	71
5.2	Trabalhos futuros	72
5.3	Contribuições	72
5.4	Trabalho submetido	73
	REFERÊNCIAS	75
	 APÊNDICES	 83
	Apêndice A – INFORMAÇÕES COMPLEMENTARES	85
A.1	Ressonância Magnética	85
A.2	Classificador SVM	87

Capítulo 1

Introdução

Neste capítulo é apresentado o contexto no qual a pesquisa está inserida, indicando o problema investigado e os objetivos a serem alcançados pela sua solução.

1.1 Contexto e motivação

A doença de Alzheimer (DA) é uma doença neurodegenerativa, progressiva, irreversível e que afeta majoritariamente a população idosa. É a forma de demência mais comum no mundo, correspondendo a até 80% de todos os casos reportados. Pelo caráter progressivo da doença e sua irreversibilidade, o diagnóstico tardio pode ser fatal. Sendo esta a terceira causa de morte após doenças do coração e câncer (ZHANG et al., 2017; ALZHEIMER'S ASSOCIATION, 2020).

As chances de desenvolvimento DA estão relacionadas a fatores genéticos e comportamentais, como hipertensão, tabagismo, histórico de diabetes, obesidade, e também ao avançado da idade, como perdas neuronais e atrofia estruturais. A Organização Mundial da Saúde (OMS) estima que a população mundial com Alzheimer aproxima-se de 35 milhões e que em 2050 irá afetar 22% das pessoas idosas (ALZHEIMER'S DISEASE INTERNATIONAL, 2020). Levando-se em consideração o cenário atual e as projeções futuras, a OMS incluiu a doença de Alzheimer como uma condição prioritária de saúde.

O diagnóstico da DA é um processo que possui diversos parâmetros e critérios, como exames clínicos e avaliações de capacidades neurológicas. A avaliação diagnóstica definitiva observa o comprometimento cognitivo da memória e outros domínios da função intelectual e, conseqüentemente, comprometimento social ou ocupacional (PRINCE et al., 2013). Como parte do exame clínico, o imageamento por ressonância magnética (RM) estrutural (FRISONI et al., 2010) desempenha um papel importante e necessário, pois permite identificar alterações em estruturas cerebrais e, portanto, evidenciar a progressão

da doença (FRISONI et al., 2010). Logo, o exame de RM funciona como uma informação complementar importante à avaliação clínica no diagnóstico precoce da DA, permitindo que a introdução de um tratamento apropriado, objetivando o retardo na progressão da DA (ALZHEIMER'S ASSOCIATION, 2020). A RM é a técnica de imagem preferida para a análise estrutural do cérebro na DA, pois fornece imagens com excelente diferenciação dos tecidos moles e boa resolução espacial, sem riscos conhecidos à saúde.

Os hipocampos são estruturas cerebrais localizadas no lobo temporal médio, que desempenham um papel importante na formação de memórias, e são bastante afetadas pela doença, apresentando alto grau de assimetria e atrofia no estágio avançado da doença. A progressão da DA também afeta as amígdalas, ventrículos e outras regiões como o córtex entorrinal. Na fase final da doença, o cérebro é afetado de forma generalizada, apresentando redução significativa do tecido cerebral (HECKERS; KONRADI, 2010; GUADALUPE et al., 2014; WOOLARD; HECKERS, 2012). A Figura 1 ilustra regiões cerebrais com significativo grau de atrofia causada pelo desenvolvimento da DA. Esse processo, que é gradual, pode ser identificado num estágio pré-clínico da doença em algumas regiões cerebrais, como o hipocampo e o lobo temporal medial. Alterações atróficas também são encontradas no estágio conhecido como comprometimento cognitivo leve (CCL), que pode ser considerado um estágio de transição entre um cérebro cognitivamente normal (CN) e a DA (COUPÉ et al., 2012; MORRIS et al., 2001).

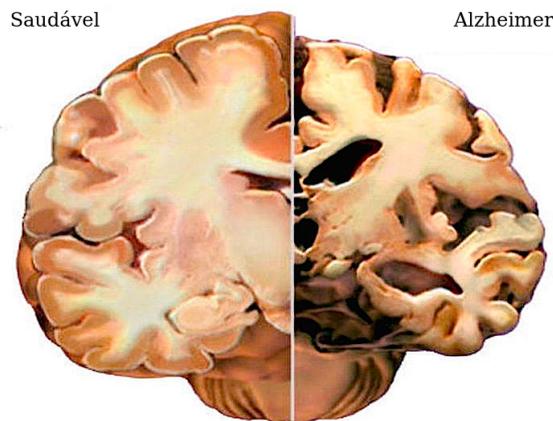


Figura 1 – Diferenças estruturais entre um cérebro saudável (cognitivamente normal) e com doença de Alzheimer. Extraído e adaptado de (ALZHEIMER'S ASSOCIATION, 2020).

O diagnóstico neuropatológico da DA depende da presença de emaranhados neurofibrilares e placas senis. O número de emaranhados neurofibrilares está intimamente ligado ao grau de demência, sugerindo que a formação de emaranhados neurofibrilares se correlaciona diretamente com a disfunção neuronal (BUTTERFIELD; HALLIWELL, 2019; MARTINS et al., 2018). No Alzheimer, o padrão regional das áreas afetadas pela formação de emaranhados neurofibrilares durante o curso da doença é relativamente conhecido e não é difícil de prever. Estudos recentes relacionados a biomarcadores da DA

(MCKHANN et al., 2011; DUBOIS et al., 2014) recomendam o uso de medidas do líquido cefalorraquidiano (CSF), Amiloide- β , tau e biomarcadores de lesão neuronal para auxiliar no diagnóstico da DA. Alguns desses biomarcadores podem ser obtidos a partir de técnicas de neuroimagem, como o imageamento por RM e tomografia por emissão de pósitrons (Positron Emission Tomography - PET). Além das informações estruturais, a formação de emaranhados neurofibrilares e o depósito de placas senis podem deixar certos padrões de valor cinza na imagem de RM que, embora possam não ser distinguíveis pelo olho humano, podem ser avaliados a partir de algoritmos de textura (SØRENSEN et al., 2016; CASTELLANO et al., 2004).

Existem vários estudos na literatura que tem por objetivo capturar os padrões sutis de intensidade do nível de cinza decorrentes da DA (ZHANG et al., 2012; LIU et al., 2017; BUSTAMAM; SARWINDA; ARDENASWARI, 2018). Os métodos comumente usados para extração de textura são o Gray Level Co-Occurrence Matrix (GLCM) (HARRALICK, 1973), o Run Length Matrix (RLM) (GALLOWAY, 1974) e o Local Binary Patterns (LBP)(OJALA; PIETIKÄINEN; HARWOOD, 1996). Esses métodos funcionam como extratores de características de textura de base estatística e estrutural, uma vez que analisam a relação espacial entre um pixel (ou voxel) e sua vizinhança (CAI et al., 2020). Outras abordagens menos comuns, mas não menos importantes para a análise de textura em AD, usam métodos baseados em transformadas, a fim de avaliar as informações da orientação de frequências espaciais de uma imagem (FENG; ZHANG; CHEN, 2020; NANNI et al., 2019).

Pesquisas recentes (HETT et al., 2021; MA; WU; KIM, 2020; MANUSCRIPT, 2018; CUI et al., 2018; HETT et al., 2018; GUO et al., 2017; JIE et al., 2016) buscam identificar alterações nos tecidos cerebrais a partir da segmentação e extração de atributos de diferentes regiões do cérebro. Estes trabalhos não se restringem à região hipocampal, sugerindo que diferentes regiões cerebrais participam do desenvolvimento da doença de Alzheimer. e que, portanto, a extração de atributos dessas regiões poderia auxiliar na obtenção de melhores resultados de classificação. Entretanto, um dos desafios que surge é de como modelar a relação de diferentes regiões segmentadas. Neste contexto, a modelagem das regiões em grafos tem sido utilizada, uma vez que grafos possuem um *framework* genérico capaz de se adaptar a diferentes problemas.

Com a modelagem de imagens de RM e suas regiões em grafos surgem outros desafios, que cada autor aborda de forma distinta, como selecionar os melhores atributos, como identificar quais conexões entre regiões são mais relevantes para classificar pacientes CN, CCL e pacientes com DA. Para imagens de RM estruturais, pesquisas recentes estão utilizando atributos de textura das regiões, como proposto pelos autores Liu et al. (2016), Liu et al. (2017), Beheshti et al. (2017) e Hett et al. (2021), sugerindo que atributos de texturas extraídos de diversas regiões podem apresentar alterações como consequência de mudanças causadas pela progressão da DA.

Outro desafio da abordagem via grafos é de como transformar um grafo em uma representação estruturada na forma de um vetor, tal que seja possível calcular similaridades por um algoritmo de aprendizado de máquina. Nesse sentido, a técnica *Graph Kernel* (GK) surgiu com o objetivo de representar grafos complexos como um vetor num espaço de baixa dimensionalidade, enquanto mantém suas propriedades. O GK é um *kernel* capaz de transformar um grafo complexo para um novo espaço, a partir de atributos em seus vértices e em suas conexões. Estudos publicados sugerem o uso de GK para classificação de imagens de doenças neurodegenerativas como o Alzheimer (SHARAEV et al., 2019; CUI et al., 2018; JIE et al., 2016; WANG; WILSON; HANCOCK, 2016; JIE et al., 2018), porém, a maioria dos trabalhos não utilizam imagens de RM estruturais, mas sim funcionais, como PET e ressonância magnética funcional (*Functional Magnetic Resonance Imaging* - fMRI). Especialmente pela facilidade técnica do uso de fMRI, dado a própria técnica já extrai atributos de ativação neuronal capazes de discriminar entre grupos de diagnóstico. Porém, para a captura das imagens de fMRI, o paciente é submetido a injeção de contraste o que pode ser danoso a saúde do paciente de DA, dado que é uma doença que afeta majoritariamente pessoas idosas. Portanto, o uso de GK com imagens de RM estruturais é uma área com bastante espaço para pesquisa, com benefícios clínicos a saúde do paciente.

1.2 Objetivos

Com base no que foi exposto, o principal objetivo deste trabalho de pesquisa foi desenvolver uma técnica computacional baseada em *graph kernel* para a classificação de grafos cerebrais construídos a partir de imagens de RM estruturais, com a extração de atributos de textura de múltiplas regiões cerebrais.

1.2.1 Objetivos específicos

- Estudar e selecionar uma base de dados que contenha imagens de RM de pacientes CN, CCL e AD.
- Preprocessar e segmentar todas as imagens de RM com uma ferramenta validada pela literatura.
- Estudar e aplicar filtros de textura para a diferenciação das imagens entre regiões e os grupos de diagnóstico.
- Desenvolver uma metodologia capaz de avaliar as características de textura das diversas regiões de interesse (ROIs) com grafos.
- Modelar os grafos a partir de regiões de interesse em conjunto de seus atributos de textura.

- Identificar os atributos de textura que conseguem separar os grupos de diagnóstico via classificação.

1.3 Organização do trabalho

Este trabalho está dividido em cinco capítulos e um apêndice, o primeiro capítulo contém a introdução ao tema de pesquisa, o seguido da revisão bibliográfica, para então apresentar a metodologia e técnicas utilizadas. Por fim, apresentamos os resultados obtidos nesta pesquisa e as conclusões. Uma breve descrição individual dos capítulos é apresentada a seguir:

- Capítulo 1: aborda as informações sobre as motivações e contexto para a pesquisa.
- Capítulo 2: contém a revisão bibliográfica do estado da arte relacionado à classificação em grafos para o auxílio diagnóstico da doença de Alzheimer.
- Capítulo 3: apresenta a metodologia (algoritmos e técnicas) e a base de dados utilizada neste trabalho.
- Capítulo 4: apresenta os resultados obtidos e discussões.
- Capítulo 5: apresenta as conclusões finais.
- Apêndice A: contém informações complementares sobre a formação de imagens de ressonância magnética e sobre o classificador SVM.

Capítulo 2

Revisão Bibliográfica

Neste capítulo é apresentado a revisão bibliográfica do estado da arte de técnicas de classificação de grafos, especialmente no uso de Graph Kernels, com foco no auxílio ao diagnóstico da doença de Alzheimer via processamento de imagens de RM.

2.1 Classificação de grafos em ressonância magnética

A modelagem de estruturas cerebrais usando grafos é amplamente utilizada na literatura de neuroimagens, especialmente com foco em doenças neurodegenerativas (JIE et al., 2016; GUO et al., 2017; CUI et al., 2018; KRIEGE; JOHANSSON; MORRIS, 2020a; HETT et al., 2021). Diversas evidências de estudos anatômicos, funcionais e psicológicos sugerem que os processos cognitivos dependem da interação distribuída de diversas regiões cerebrais (SPORNS, 2014; WEE et al., 2014). Essas interações podem ser caracterizadas por um grafo cerebral, quantificando suas propriedades pela teoria dos Grafos.

Porém, visto que grafos não tem uma representação direta em um espaço de atributos, como vetores, isso dificulta o seu uso em classificação de imagens, o que leva ao problema de como calcular a similaridade entre grafos. Diversas propostas foram apresentadas para a extração de dados e atributos de grafos com aplicação em classificação de imagens, em especial citamos o *Graph Kernel*, que é um método que pode ser utilizado em máquinas de vetores suporte (*Support Vector Machines* - SVM), facilitando, assim, a comparação entre grafos (RICHIARDI et al., 2013a).

2.1.1 Trabalhos correlatos com *Graph Kernel*

Em 2018 Cui et al. (2018) propuseram uma solução para trabalhar com grafos cerebrais que envolve selecionar os nós e arestas dos grafos de maneira a obter bons resultados de

classificação. Para isto, utilizaram um *framework* de classificação baseado nas árvores geradoras de custo mínimo (*minimum spanning tree* - MST). A MST foi inicialmente aplicada em grafos cerebrais por Lee, Kim e Jung (2006) e mais recentemente no trabalho de Guo et al. (2017) que utilizou tal abordagem para classificação da DA.

O imageamento por fMRI permite extrair a ativação de regiões de interesse na forma de séries em diversos pontos do tempo (séries temporais). Cui et al. (2018) utilizaram neste trabalho imagens de fMRI da base de dados Alzheimer's Disease Neuroimaging Initiative (ADNI) e o trabalho incluiu 68 pacientes, divididos da seguinte maneira: 21 pacientes com DA, 25 com CCL e 22 CN. As imagens foram pré-processadas e segmentadas em 90 ROIs com o algoritmo Automated Anatomical Labeling (AAL) (TZOURIO-MAZOYER et al., 2002). O grafo foi construído utilizando cada ROI como um nó e o peso de suas conexões (arestas do grafo) foi definido pelo coeficiente de correlação de *Pearson* entre as ROIs formando cada aresta, gerando, assim, um grafo completamente conectado. A partir deste grafo e aplicando o algoritmo de Kruskal (KRUSKAL, 1956), os autores encontraram a MST e, com o algoritmo gSpan (YAN; HAN, 2002) capturaram os subgrafos mais frequentes dentro de um grupo. Dessa maneira, os autores puderam identificar quais subgrafos foram os mais discriminantes entre os grupos de diagnóstico.

Assim, os grafos são construídos a partir dos subgrafos mais frequentes, preservando sua topologia de classe e mantendo a discriminabilidade. Aplicando o *Graph Kernel* de Weisfeiler-Lehman (SHERVASHIDZE et al., 2011), os autores computam a matriz de similaridades entre todos os grafos, par a par, para então, com a matriz de *Kernel*, aplicar *Kernel Principle components Analysis* (kPCA) (SCHÖLKOPF; SMOLA; MÜLLER, 1997) com os 8 maiores autovalores. Dessa maneira, os autores evitam o uso de métodos para remoção de arestas baseados em limiar que, de acordo com os autores, podem afetar negativamente o desempenho de classificação. Finalmente, a classificação é realizada usando um classificador SVM com 100 repetições de 10-fold de validação cruzada.

Os resultados obtidos pelo trabalho superam os resultados de métodos por limiar. Para o caso CCL×DA os autores obtiveram 77,3% de acurácia, com 100% de especificidade e 54,1% de sensibilidade, e AUC (*Area Under Curve* - AUC) de 0,97. Para o CN×CCL os valores obtidos foram 98,3% de acurácia, 100% de especificidade e 96,6% de sensibilidade, e AUC de 0,99. Por fim, para a classificação CN×DA os valores foram de 91,3% de acurácia, 100% de especificidade e 82,5% de sensibilidade, e AUC de valor 1.

Ainda neste trabalho de Cui et al. (2018) também apresentam um comparativo do método desenvolvido com métodos baseados em limiar (*Threshold-based*), onde definindo um limiar de esparsidade de 40%, obtiveram valores de acurácia de 63,3%, 87,5% e 65,8% para CN×CCL, CN×DA e CCL×DA, respectivamente. O trabalho apresenta as seguintes limitações: (a) baixo valor de sensibilidade quando comparado às outras métricas apresentadas e (b) uso de um número pequeno de imagens por grupo, o que pode ter enviesado os resultados.

O trabalho apresentado por Wang, Wilson e Hancock (2016) construiu um *Graph Kernel* com base na divergência de Jensen-Shannon utilizando o já existente Jensen-Shannon *kernel* (MARTINS et al., 2009), o qual utiliza a entropia de Bose-Einstein e aplica a técnica kPCA para mapear grafos em um espaço de características de baixa dimensionalidade, utilizaram os três primeiros autovetores calculados a partir da matriz de Jensen-Shannon. Os autores utilizaram o método proposto para a classificação de imagens fMRI no auxílio diagnóstico da doença de Alzheimer.

A base de dados utilizada foi composta por imagens de fMRI da ADNI contendo 30 imagens de AD, 34 de Comprometimento cognitivo leve tardio (CCLT), 47 de Comprometimento cognitivo leve precoce (CCLP) e 38 de CN. Os autores construíram os grafos utilizando como vértice cada ROI e a partir dos grafos completos removeram as arestas tal que o coeficiente de correlação-cruzada entre as séries temporais fosse menor que 40% de todos os coeficientes. Devido o número de imagens de cada classe ser baixo, os autores utilizaram o método *leave-one-out cross-validation* para usar todos os grafos como dados de teste na classificação pela técnica de análise discriminante de Fisher. Com o *Graph Kernel* proposto, os autores obtiveram valores de acurácia de 93,33% para a DA, 100% para a classe CCLT, 89,36% para a CCLP e 92,11% para a CN.

Apesar dos altos valores de acurácia, os autores não apresentaram a estratificação dos dados, não indicaram o uso de pesos nas arestas, e não apresentaram valores de desvio padrão da acurácia de classificação. Devido a baixa quantidade de imagens por classe e a falta de indicadores como a sensibilidade e especificidade, podem dificultar o entendimento dos resultados apresentados.

Em 2016 Jie et al. (2016) apresentaram um trabalho em que os autores desenvolveram um *Graph Kernel* específico para a análise de imagens de cérebro que difere dos outros *Graph Kernels* propostos na literatura. O método desenvolvido pelos autores considera cada nó como único em cada grafo e entre os grafos. Além disso, o método percorre a vizinhança de cada nó, buscando por similaridades topológicas entre os grafos e levando em consideração propriedades tipicamente vistas em grafos com caminhos mais curtos entre nós, triângulos e isomorfismo. Dessa maneira, o método proposto produz uma série de sub-grafos para cada nó, assim, o *kernel* desenvolvido consegue computar a similaridade entre sub-grafos de nós correspondentes.

Para validar o método proposto, os autores utilizaram a base de dados ADNI com imagens fMRI e RM de 149 pacientes, dentre as quais 50 são da classe CN e 99 da classe CCL, sendo ainda o grupo CCL composto por dois subgrupos de pacientes, 56 CCLP e 43 CCLT. As imagens de fMRI de cada paciente possuíam 140 leituras funcionais, isto significa que cada paciente possui 140 valores de ativação neuronal por voxel. As imagens fMRI foram preprocessadas e alinhadas com o primeiro volume usando o software FSL FEAT¹, que possui ferramentas de análise específicas para imagens fMRI. Em conjunto, os autores

¹ <http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FEAT>

utilizaram imagens T1-ponderadas e o pré-processamento incluiu as seguintes etapas: (a) remoção do crânio, (b) segmentação das imagens nas classes substâncias cinzenta (*gray matter* - GM), branca (*white matter* - WM) e líquido cefalorraquidiano (*cerebral spinal fluid* - CSF) utilizando o FSL FAST² e (c) correção das imagens T1-ponderadas com o primeiro volume das imagens fMRI. Finalmente, com as imagens T1-ponderadas e fMRI alinhadas, os autores aplicaram o algoritmo AAL proposto por Tzourio-Mazoyer et al. (2002) para segmentar as imagens T1-ponderadas em 90 regiões de interesse (45 regiões por hemisfério).

Para a criação dos grafos cerebrais, cada nó foi definido por uma região segmentada e o coeficiente de correlação de Pearson foi usado como peso das arestas. Por fim, a normalização dos coeficientes das arestas foi realizada pela transformação Z de Fisher (FISHER, 1915). Partindo de um grafo completo, em que cada nó é conectado a todos os demais nós, os autores extraíram os coeficientes de aglomeração local (RUBINOV; SPORNS, 2010) para cada grafo construído, calculando o *p*-valor de cada vetor usando o método *t-test*, e usaram o *p*-valor para remover os atributos abaixo de um certo limiar. Os autores definiram vários limiares com o intuito de representar topologias distintas entre os grafos, favorecendo o *kernel* desenvolvido.

A classificação foi feita com o SVM utilizando o *kernel* desenvolvido, com parâmetro *C* igual a um, e cinco valores distintos de limiares para a seleção de atributos do grafo. O método proposto resultou em 82,6% de acurácia para o caso CN×CCL, e 74,8% de acurácia para o caso CCLP×CCLT, enquanto que os melhores valores de acurácia correspondentes para outros *kernels* da literatura foram de 76,5% e 70,7% respectivamente. Os autores também reportaram que, ao combinar diferentes valores de limiares para a seleção de atributos do grafo, tanto a acurácia quanto a AUC responderam positivamente, porém, os autores não indicaram a quantidade de arestas ou regiões removidas durante a limiarização.

2.1.2 Trabalhos correlatos sem *Graph Kernel*

Para elaborar uma representação discriminante das estruturas anatômicas tal que separe pacientes em grupos diagnósticos, Liu et al. (2016) propuseram um grafo hierárquico abrangendo a imagem de todo o cérebro dos pacientes. Para isso, a imagem foi dividida em 90 ROIs, consideradas em 4 diferentes níveis de abstração. Para tanto, os autores utilizaram 710 imagens de RM T1-ponderadas da base de dados ADNI, sendo estas divididas em 230 de pacientes CN, 280 com CCL e 200 de pacientes com DA. O grupo CCL foi composto por 120 imagens de pacientes que converteram para a DA num período de 18 meses (definido como subgrupo CCLpr), e 160 imagens de pacientes CCL que não desenvolveram DA no mesmo período (CCLe). Nas análises, as imagens foram estratificadas por idade dos pacientes.

² <http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FAST>

No pré-processamento das imagens, os autores utilizaram o algoritmo *Non-parametric Non-uniform intensity Normalization* (N3) (SLED; ZIJDENBOS; EVANS, 1998) para a correção da não-homogeneidades de intensidade, a ferramenta BET (SMITH, 2002) para a remoção do crânio e cerebelo, o pacote FAST (ZHANG; BRADY; SMITH, 2001) para a segmentação dos tecidos GM, WM e CSF e, por fim, o algoritmo FLIRT (JENKINSON; SMITH, 2001; JENKINSON et al., 2002), com transformações afins, para o alinhamento espacial de todas as imagens.

Para a obtenção das 90 ROIs (45 por hemisfério) foi utilizado o algoritmo AAL proposto por Tzourio-Mazoyer et al. (2002). O grafo foi montado hierarquicamente em quatro níveis, sendo a primeira camada do grafo representada pelas 90 regiões segmentadas e as demais camadas pela concatenação de ROIs com base numa regra desenvolvida pelos autores. A quarta e última camada, contém somente uma ROI, que é o próprio cérebro (concatenação de todas as ROI's).

Seis atributos de textura foram extraídos de cada ROI utilizando a técnica *Gray Level Co-occurrence Matrix* (GLCM). Os autores utilizaram a média de cada atributo de textura de uma ROI para formar seu respectivo vetor de seis atributos. A seleção de atributos foi realizada com uma equação de F-score adaptada pelos autores, que ranqueia os atributos mais discriminativos. As arestas do grafo foram avaliadas entre todas as possíveis combinações de pares de ROIs conforme a correlação de Pearson, para seleção das arestas que melhor representam o grafo de cada paciente. Assim, os autores construíram um vetor de atributos de conectividade por imagem com 12.561 dimensões.

Os autores utilizaram o classificador MKBoost (XIA; HOI, 2012) com validação cruzada de 10 *folds* e obtiveram 94,65% de acurácia, 0,954 de AUC, 95,03% de sensibilidade e 91,76% de especificidade para a classificação CN×DA. Para a classificação CCL×DA, os resultados obtidos foram de 89,63% de acurácia, 0,907 de AUC, 91,55% de sensibilidade e 86,25% de especificidade. Ainda, para a classificação CN×CCL os resultados foram de 85,79% de acurácia, com 0,826 de AUC, com 88,91% de sensibilidade e 80,34% de especificidade. Finalmente, para a classificação CCLpr×CCLe, os resultados obtidos foram de 72,08% de acurácia, 0,716 de AUC, 75,11% de sensibilidade e 71,05% de especificidade.

Com base nos resultados acima, os autores tentaram mostrar o potencial uso de atributos de textura GLCM extraídos de regiões segmentadas na classificação das imagens de RM. No ano seguinte a esse trabalho, Liu et al. (2017) adicionaram os atributos de cada ROI para a classificação, isto é, os nós do grafo. Nessa nova proposta, os autores utilizaram um vetor de atributos de textura com seis dimensões como descritor para cada ROI e mantiveram o mesmo método para compor as arestas. De forma similar ao trabalho anterior, o algoritmo F-Score foi aplicado em cada nó do grafo para a seleção de atributos. A classificação foi realizada por um *Ensemble* de classificadores MKoost, um para cada um dos seis atributos dos nós. Além disso, os autores utilizaram um classificador MKBoost aplicado somente aos atributos de arestas.

Os resultados das mudanças propostas no trabalho de Liu et al. (2016) resultaram nos seguintes resultados: 95,37% de acurácia, 0,967 de AUC, 95,49% de sensibilidade e 96,08% de especificidade para a classificação de CN×DA. Já na classificação CCL×DA, os valores foram 90,41% de acurácia, 0,919 de AUC, 92,83% de sensibilidade e 88,82% de especificidade. Por fim, na classificação CN×CCL, os resultados obtidos pelo trabalho foram de 86,56% de acurácia, 0,854 de AUC, 90,74% de sensibilidade e 84,83% de especificidade.

Em 2018 Hett et al. (2018) propuseram um *framework* baseado em grafos visando capturar similaridades inter-pacientes e variabilidades intra-paciente para melhor detecção da DA. Para capturar as similaridades, eles utilizaram um método baseado em *patches* de pontuação de voxels, desenvolvido por Coupé et al. (2012), enquanto que, para capturar a variabilidade entre as estruturas cerebrais, os autores utilizaram uma modelagem por grafo.

A base de dados utilizada no trabalho foi a ADNI, com 683 imagens T1-ponderadas, divididas em quatro classes, sendo 228 CN, 100 CCLpr, 164 CCLe e 191 DA. O grupo CCLpr possui pacientes que desenvolveram DA no decorrer de 36 meses da primeira aquisição das imagens. As imagens foram estratificadas pela idade dos pacientes, com média de 75 anos e com um desvio padrão de 7,4 anos. O pré processamento envolveu a redução de ruídos das imagens, usando o método *adaptive non-local means filter* (MANJÓN et al., 2010), a correção do *bias field* com o algoritmo N4ITK (TUSTISON et al., 2010), a padronização de intensidade, a segmentação de regiões com o método *non-local label fusion* (GIRAUD et al., 2016) e por fim correção automática de erros de segmentação com o algoritmo proposto por Wang et al. (2011).

Para capturar as alterações causadas pela DA, o método de pontuação Patch-Based Grading (PBG) de Coupé et al. (2012) define um valor entre -1 e 1 para cada voxel, conforme o grau de alteração do mesmo pela doença, sendo -1 quando o voxel é definido como pertencente a um paciente DA e 1 para um paciente CN, levando em consideração a alteração estrutural da vizinhança de cada voxel. Assim, a técnica PBG atribui um valor entre -1 e 1 à cada voxel da imagem.

A segmentação aplicada no pré-processamento dividiu a imagem em 134 regiões. Cada região foi utilizada como um nó do grafo e, como cada região possui valores entre -1 e 1 obtidos do algoritmo PBG, os autores utilizaram um histograma para representar os valores em cada ROI. Cada histograma teve seu número de *bins* calculado pela regra de Sturge (STURGES, 1926). Como atributo do nó, os autores utilizaram a média do histograma e para o peso das arestas utilizaram uma equação que leva em consideração a distância de Wasserstein (RUBNER; TOMASI; GUIBAS, 2000) e desvio padrão. Por fim, os atributos foram normalizados pelo método z-score.

Para a seleção de arestas e nós discriminantes, os autores aplicaram o método *Elastic Net regression* (EN) para obter uma representação esparsa das arestas e nós mais discrimi-

minativos. A classificação foi realizada usando os classificadores SVM de *kernel* linear e Random Forest (RF). A seleção de atributos e o treinamento dos classificadores foram feitas com os grupos CN e DA. Usando o classificador RF para o grupo CCLpr×CCLe, os resultados obtidos foram 76,5% de acurácia, 81,7% de sensibilidade e 68% de especificidade. Já para o classificador SVM, os resultados foram 75,8% de acurácia, 82,3% de sensibilidade e 65% de especificidade. Os autores não indicaram quais regiões foram mais relevantes para o grafo, nem quantos nós e arestas foram removidos para obter seus resultados.

Recentemente Hett et al. (2021) propuseram uma continuação de seu trabalho anterior (HETT et al., 2018), expandindo-o para outras regiões cerebrais e com adição de novos atributos para classificação. Neste trabalho, os autores utilizaram 559 imagens da base de dados ADNI, divididas em 213 CN, 126 CCLpr, 90 CCLe e 130 DA e os mesmos algoritmos de pré-processamento do trabalho anterior, alinhando as imagens ao espaço MNI. As imagens foram segmentadas em 133 regiões com um método de segmentação multiescala (GIRAUD et al., 2016) e, em paralelo, os autores também segmentaram os hipocampos em cinco sub-regiões.

A extração de atributos por voxel em cada região foi realizada pelo método PBG utilizado no trabalho anterior. Esse método define um valor de -1 a 1 a cada voxel das imagens, conforme sua probabilidade de ter sofrido alterações relacionadas a DA. Os grafos foram construídos utilizando cada região como um vértice, e seus atributos foram computados a partir dos valores dos voxels da região definidos pela técnica PBG. Os atributos dos vértices foram definidos pela estimação de uma distribuição de probabilidade da região a partir de um histograma. O histograma teve seu número de bins calculado a partir do número médio de voxels dentre todas as regiões segmentadas com a regra de Sturge. As arestas dos grafos foram calculadas pela distância de Wasserstein entre as regiões adjacentes. Como novidade deste trabalho, e utilizando a mesma metodologia de construção de grafos, os autores desenvolveram grafos a partir das cinco sub-regiões dos hipocampos. A seleção de vértices e arestas mais discriminantes em ambos os grafos foi feita utilizando o método EN.

Os autores também propuseram como novidade neste trabalho, a adição do resultado de cinco testes cognitivos aplicados para cada paciente. Assim, a etapa de classificação foi feita em uma cascata de classificadores RF, onde o grafo do cérebro e o grafo das sub-regiões dos hipocampos foram classificados independentemente. Seus resultados então foram utilizados em conjunto com os testes cognitivos para um classificador RF final. Os atributos dos grafos e dos testes cognitivos foram normalizados utilizando o método z-score. Os resultados de classificação para o caso CCLpr×CCLe foram de 0.85 de AUC, 80.6% de acurácia, 87.3% de sensibilidade e 71.1% de especificidade. Os autores também realizaram experimentos para a classificação CN×DA, onde obtiveram 91.6% de acurácia. Os resultados das classificações sugerem que a inclusão das sub-regiões dos hipocampos

resultou em uma melhora na classificação. Os autores também ressaltam a importância dos valores dos testes cognitivos na classificação, em que a classificação de CCLpr×CCLe, sem a adição dos testes cognitivos resultou em 76% de acurácia, um resultado inferior ao obtido pela união dos testes cognitivos.

A Tabela 1 apresenta um sumário dos trabalhos descritos neste capítulo.

Trabalho	Base de dados	Técnicas	Algoritmo de Classificação	Resultados
				Classes: Acurácia (%)
Cui et al. (2018)	ADNI, fMRI	ROIs, SA, MST, GK	SVM	CCL×DA: 77,3 CN×CCL: 98,3 CN×DA: 91,3
Wang, Wilson e Hancock (2016)	ADNI, fMRI	ROIs, Limiar, GK	Análise discriminante de Fisher	DA: 93,33 CCLT: 100 CCLP: 89,36 CN: 92,11
Jie et al. (2016)	ADNI, fMRI	ROIs, Limiar, GK	SVM	CN×CCL: 82,6 CCLP×CCLT: 74,8
Liu et al. (2016)	ADNI, MRI	ROIs, SA, GLCM	MKBoost	CCL×DA: 89,63 CN×CCL: 85,79 CN×DA: 94,65
Liu et al. (2017)	ADNI, MRI	ROIs, SA, GLCM	Ensemble de MKBoost	CCL×DA: 90,41 CN×CCL: 86,6 CN×DA: 95,37
Hett et al. (2018)	ADNI, MRI	ROIs, Histogramas, SA	SVM, Random Forest	CCLpr×CCLe: 76,5
Hett et al. (2021)	ADNI, MRI	ROIs, Histogramas, SA	Random Forest	CCLpr×CCLe: 80,6

SA: Seleção de Atributos

Tabela 1 – Resumo dos trabalhos recentes relacionados ao tema dessa pesquisa.

2.1.3 Discussão e considerações finais

Este capítulo apresentou alguns dos principais trabalhos propostos na literatura nos últimos anos para classificação de grafos construídos com atributos de imagens para o auxílio diagnóstico da doença de Alzheimer. Analisando os trabalhos apresentados, nota-se a heterogeneidade no uso de grafos, como por exemplo o uso de diversas metodologias para a seleção de arestas. Cui et al. (2018), por exemplo, propôs a aplicação do algoritmo MST em todos os grafos para identificar a rota de arestas que mais se repete dentro de um grupo, enquanto os outros trabalhos, quando utilizaram as arestas na análise, aplicaram métodos baseados em limiar para remover as arestas não discriminantes entre as classes. Em relação à classificação dos grafos, Cui et al. (2018), Wang, Wilson e Hancock (2016) e Jie et al. (2016) utilizaram *graph kernel* como representação dos grafos e seus atributos em um espaço de menor dimensionalidade. Ao utilizar *graph kernel*, os autores puderam focar o desenvolvimento e análise nas etapas anteriores à classificação,

como as de seleção de atributos, normalizações e de definição da estrutura dos grafos. A partir dos trabalhos apresentados foi possível notar a carência de pesquisas utilizando algoritmos de textura como o RLM e LBP. Em específico, o nosso trabalho é possivelmente o único a utilizar a combinação de *graph kernel* com atributos de textura em imagens de ressonância magnética estrutural.

Capítulo 3

Metodologia

Neste capítulo é apresentado em detalhes cada etapa do projeto apresentado neste trabalho de pesquisa, como a base de dados, os algoritmos de textura, métricas e a metodologia da construção dos grafos.

3.1 Visão geral

Essa pesquisa consistiu no desenvolvimento de uma técnica para a classificação de grafos, construídos de informações extraídas de diversas regiões cerebrais de imagens de RM, em três grupos de diagnóstico, CN, CCL e DA, visando o auxílio ao diagnóstico da doença de Alzheimer. Para isso, uma técnica de segmentação automática foi usada para obter as diversas regiões de interesse das quais foram extraídas as distribuições de probabilidade de 22 atributos de textura para compor os atributos dos grafos. Visando obter resultados menos enviesados e sem a necessidade de reprocessamento, separamos a base de dados em dois grupos, um exclusivo para a classificação e outro composto somente de imagens de CN e DA que foi utilizada para a estimação de parâmetros necessários ao projeto.

A Figura 2 representa esquematicamente a abordagem utilizada para a classificação das imagens neste projeto. Partindo de uma base de dados com imagens de RM, primeiro realiza-se o pré-processamento, a segmentação das imagens em diversas regiões de interesse. A seguir, utiliza-se três algoritmos para a extração de textura das imagens pré-processadas e as máscaras das regiões de interesse, obtidas no estágio anterior, são utilizadas para obter uma matriz de características de textura por paciente. A Figura 2 também demonstra os passos utilizados com o grupo de estimação de parâmetros. Cada etapa do fluxograma é detalhada nas seções subsequentes deste capítulo.

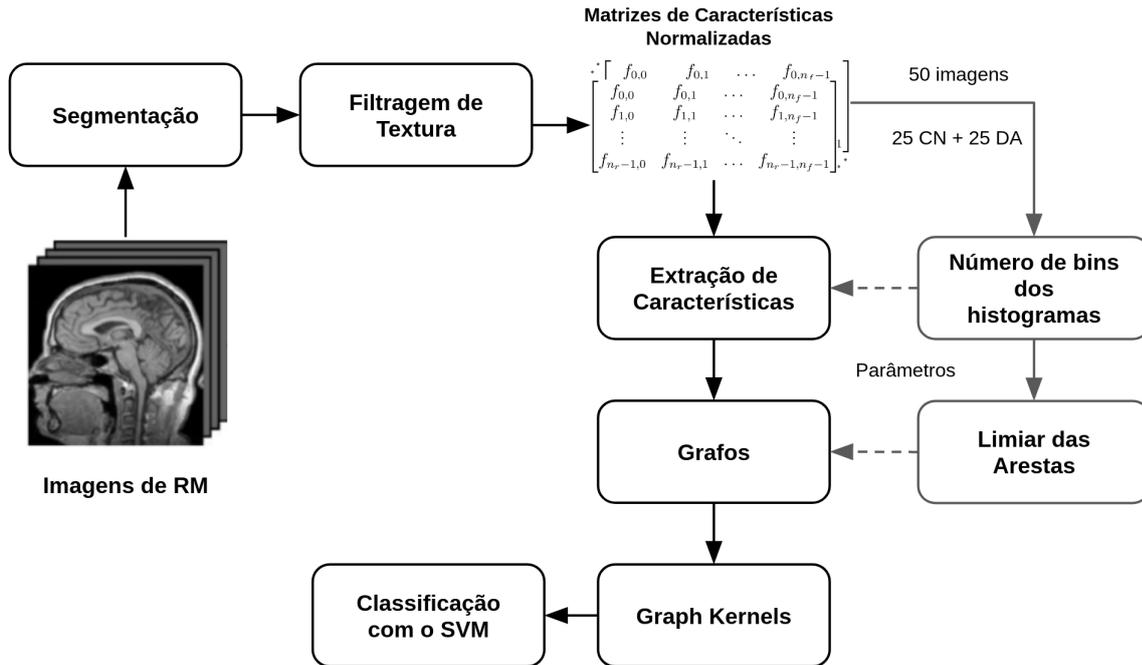


Figura 2 – Diagrama de blocos indicando o fluxo de funcionamento do projeto.

3.2 Base de dados

A base de dados *Alzheimer's Disease Neuroimaging Initiative* (ADNI) (JR et al., 2008) é uma iniciativa lançada em outubro de 2004 por parcerias público-privadas entre mais de 20 empresas e instituições públicas que tem por objetivo a busca por maneiras de identificar e acompanhar o desenvolvimento da doença de Alzheimer, elencando novos biomarcadores que possam auxiliar no desenvolvimento de novos métodos de diagnóstico precoce. Para este fim, eles provêm sem restrições as imagens coletadas pela iniciativa para os cientistas interessados, o que pode ser feito via página do *USC's Laboratory of Neuro Imaging's* (LONI) na área *Image & Data Archive* (IDA)¹. A base de dados contém imagens RM estruturais e funcionais, *Positron Emission Tomography* (PET), *Diffusion Tensor Imaging* (DTI), além de informações genéticas, bioquímicas e pontuações em testes de cognição.

Para este estudo foi utilizado um total de 474 imagens de RM estrutural de indivíduos na faixa etária entre 68 e 81 anos, divididos em três grupos de diagnóstico: CN, CCL e DA. As imagens foram obtidas usando *scanners* de 1.5T e 3T das três principais fabricantes (Philips, General Electric e Siemens) com a sequência *Magnetization Prepared Rapid Gradient Echo* (MPRAGE)². A Tabela 2 apresenta os dados demográficos dos indivíduos utilizados neste estudo. Deste total de imagens, 50 imagens (25 CN e 25 DA) foram selecionadas aleatoriamente para a estimação de parâmetros dos métodos desenvolvidos.

¹ <https://ida.loni.usc.edu/login.jsp?project=ADNI>

² "These MPRAGE files are considered the best in the quality ratings and have undergone gradwarping, intensity correction, and have been scaled for gradient drift using the phantom data." - <http://adni.loni.usc.edu/methods/mri-tool/mri-analysis/>

	CN	CCL	DA
No. pacientes	200	153	121
Idade	73.7 ± 4.2	76.6 ± 4.2	76.9 ± 4.2
MMSE	29.5 ± 0.4	26.9 ± 0.8	21.9 ± 2.6
Gênero (F:M)	101 : 99	57 : 96	55 : 66

Nota: *Mini-Mental State Examination* (MMSE)

Tabela 2 – Conjunto de imagens da base de dados ADNI.

3.3 Preprocessamento e segmentação

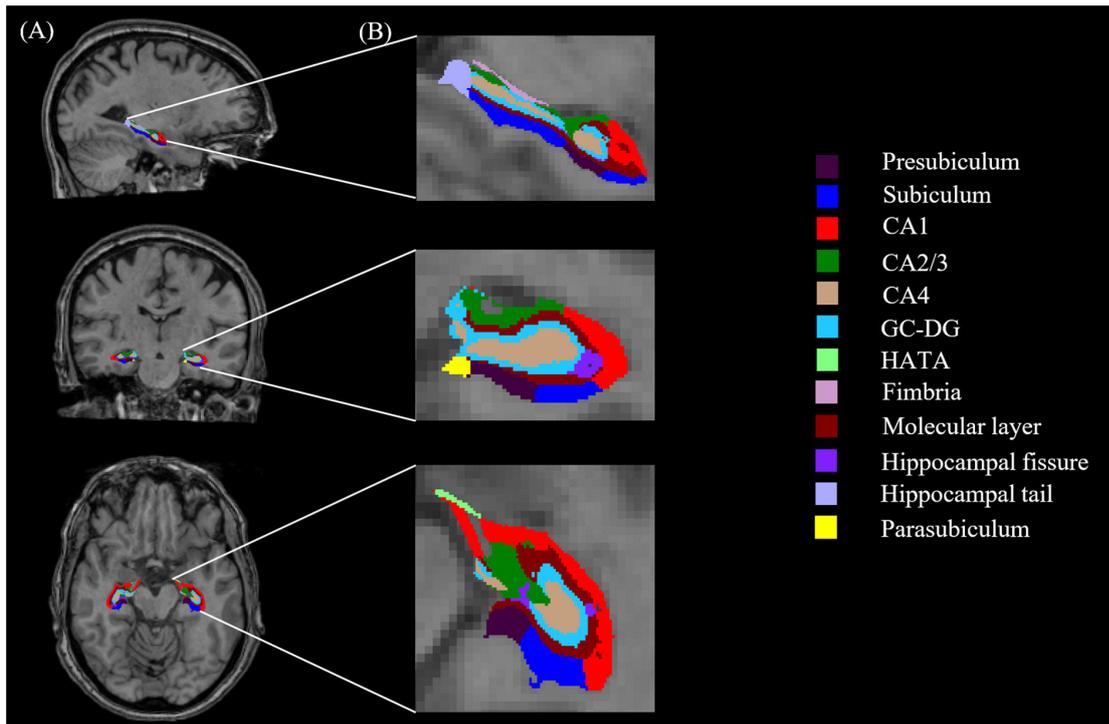
3.3.1 FreeSurfer

FreeSurfer³ é um conjunto de programas para o processamento e análise de imagens médicas amplamente utilizado pela comunidade médica, que permite a segmentação automática de diversas regiões corticais e subcorticais do cérebro em imagens de RM. Nesse projeto, o FreeSurfer foi escolhido por proporcionar uma segmentação robusta (YAAKUB et al., 2020) à atrofia estruturais e por disponibilizar algoritmos de pré-processamento comumente utilizados para a análise de imagens neurais, como remoção do crânio, correção de intensidade (TUSTISON et al., 2010), entre outros. Em sua versão mais recente, o FreeSurfer adiciona um algoritmo proposto por Iglesias et al. (2015) que segmenta o hipocampo, que é uma região importante no desenvolvimento da doença de Alzheimer, em 12 sub regiões, conforme ilustra a Figura 3. Estudos recentes (ZHAO et al., 2019; ZHENG et al., 2018; HANSEN et al., 2021) utilizaram este novo algoritmo para auxiliar no diagnóstico da doença da Alzheimer, utilizando somente o volume de cada região segmentada. Como por exemplo o estudo de Zhao et al. (2019) que mostrou que as regiões CA1, subiculum, presubiculum, molecular layer e fimbria reduziram significativamente os seus volumes na progressão de CCL para DA. O trabalho proposto por (ZHENG et al., 2018) sugere que a atrofia de sub regiões do hipocampo tem relação com o declínio da pontuação dos pacientes em exames cognitivos, como o Mini-Mental State Examination (MMSE).

3.3.2 Métodos utilizados

Todas as imagens neste projeto foram processadas utilizando o pipeline principal do FreeSurfer na versão 7.1.1, que possui 31 etapas, incluindo correção de movimento, remoção de *bias-field*, remoção do crânio, normalização de intensidade, transformação automática Talairach (COLLINS et al., 1994), tesselação de WM e GM, segmentação dos tecidos WM e GM de diversas regiões, incluindo hipocampo, amígdala e ventrículos. (FISCHL et al., 2004; SEGONNE et al., 2004; DESIKAN et al., 2006; TUSTISON et al., 2010)

³ <http://surfer.nmr.mgh.harvard.edu>



Fonte: Zheng et al. (2018)

Figura 3 – Segmentação de sub regiões do hipocampo pelo algoritmo proposto por Iglesias et al. (2015).

Ao fim da execução do pipeline do FreeSurfer é possível extrair informações de cada região, como o volume de GM, sendo este um biomarcador relevante para o diagnóstico da doença de Alzheimer pois permite evidenciar a atrofia de regiões.

A segmentação padrão do FreeSurfer utiliza o método proposto por Fischl et al. (2002), que segmenta o cérebro em 43 regiões, dentre elas 16 regiões por hemisfério usando o atlas aseg proposto por Destrieux et al. (2010). Além do atlas aseg, neste projeto também foi utilizado o atlas DKT40 (DESIKAN et al., 2006; KLEIN; TOURVILLE, 2012; FISCHL et al., 2002), que divide um cérebro em 62 regiões (31 regiões por hemisfério). Por se tratar de um cérebro adulto, o atlas DKT40 tem sido especialmente utilizado pela comunidade médica em trabalhos envolvendo a segmentação de imagens de pacientes com DA. Yaakub et al. (2020) comparou diversos atlas, incluindo o DKT, para avaliar suas capacidades de segmentação e, assim, auxiliar os pesquisadores a escolher o melhor atlas para as suas pesquisas. Seus resultados sugerem que dentre diversos atlas, o DKT do FreeSurfer desempenha melhor em regiões mais propícias as atrofias causadas pela DA, como a do córtex entorhinal, e também sugere que sua segmentação se torna mais robusta à dilatação dos ventrículos e, portanto, a seleção desse atlas é uma ótima escolha para a segmentação de imagens RM de pacientes com DA.

As imagens extraídas da ADNI, que possuem a resolução de $170 \times 256 \times 256$ voxels, são reajustadas para a resolução $256 \times 256 \times 256$, com objetivo de manter o padrão e o alinhamento ao espaço do FreeSurfer. Após a segmentação, cada ROI é extraída como

Índice	Regiões - Aseg	Índice	Regiões - Aseg	Índice	Regiões - Aseg
0	Cerebral-White-Matter	9	Pallidum	18	Optic-Chiasm
1	Cerebral-Cortex	10	Hippocampus	19	CC_Posterior
2	Lateral-Ventricle	11	Amygdala	20	CC_Mid_Posterior
3	Inf-Lat-Vent	12	Accumbens-area	21	CC_Central
4	Cerebellum-White-Matter	13	VentralDC	22	CC_Mid_Anterior
5	Cerebellum-Cortex	14	Vessel	23	CC_Anterior
6	Thalamus	15	choroid-plexus	24	3rd-Ventricle
7	Caudate	16	CSF	25	4th-Ventricle
8	Putamen	17	WM-hypointensities	26	Brain-Stem

Nota: Regiões de índice de 0 à 15 são divididas entre hemisfério esquerdo e direito.

Tabela 3 – Regiões segmentadas pelo atlas padrão do FreeSurfer.

Índice	Regiões - DKT	Índice	Regiões - DKT
0	caudalanteriorcingulate	16	parsorbitalis
1	caudalmiddlefrontal	17	parstriangularis
2	cuneus	18	pericalcarine
3	entorhinal	19	postcentral
4	fusiform	20	posteriorcingulate
5	inferiorparietal	21	precentral
6	inferiortemporal	22	precuneus
7	isthmuscingulate	23	rostralanteriorcingulate
8	lateraloccipital	24	rostralmiddlefrontal
9	lateralorbitofrontal	25	superiorfrontal
10	lingual	26	superiorparietal
11	medialorbitofrontal	27	superiortemporal
12	middletemporal	28	supramarginal
13	parahippocampal	29	transversetemporal
14	paracentral	30	insula
15	parsopercularis		

Nota: As regiões segmentadas por este atlas possuem representantes em ambos os hemisférios, esquerdo e direito.

Tabela 4 – Regiões segmentadas pelo atlas DKT40 do FreeSurfer.

uma máscara binária alinhada às imagens processadas pelo FreeSurfer. Todas as ROIs segmentadas pelo FreeSurfer usando o método proposto por Fischl et al. (2002) estão nomeadas na Tabela 3 e na Tabela 4 para o atlas DKT. Com a segmentação dos dois atlas, para cada imagem extraímos 105 regiões de interesse.

As imagens de RM fornecidas pela ADNI não incluem somente o cérebro, mas também a imagem completa da cabeça, que inclui cérebro, caixa craniana e o pescoço do paciente. Entretanto, como a caixa craniana e o pescoço são estruturas não aproveitadas na análise, usando o FreeSurfer, na etapa conhecida como remoção de crânio (*skull stripping*), obtivemos uma máscara binária somente do cérebro.

Como procedimento complementar, avaliamos todas as regiões segmentadas de nosso banco de imagens e removemos as regiões com números de voxels fora de um intervalo definido $[N_{min}, N_{max}]$, onde N_{min} é o número médio de voxels do córtex entorrinal e N_{max} um valor definido como 25.000 voxels. Este procedimento reduziu o número de regiões

de 105 para 92. Este critério foi definido a partir de análises gráficas da distribuição de valores médios por região entre os pacientes, e assim foi usado para evitar a inclusão de estruturas cerebrais muito pequenas e muito grandes, que poderiam afetar nossa análise. O que foi possível dado que todas as imagens estão todas alinhadas a um mesmo espaço estereotáxico.

3.4 Filtros de textura

O conceito de textura numa imagem digital em níveis de cinza pode ser definido como um padrão ou distribuição de valores de intensidades de cinza em uma certa região de interesse (CASTELLANO et al., 2004). A análise via textura pode representar características de regiões de uma imagem e assim servir para a discriminação de uma dada região quando comparada a outras regiões da mesma imagem. A análise de textura permite a extração de atributos de uma imagem médica, como nos trabalhos de Liu et al. (2017) e Liu et al. (2016), que mostraram a capacidade discriminativa de atributos de textura para a classificação da doença de Alzheimer.

Nesta pesquisa foram utilizadas três técnicas para a extração de atributos de textura das imagens da base de dados ADNI, as técnicas são: a *Gray Level Co-Occurrence Matrix* (GLCM), a *Run Length Matrix* (RLM) e a *Local Binary Patterns* (LBP). Com as técnicas citadas, extraímos 22 mapas de textura de cada indivíduo. Todos os filtros foram executados na área definida pela máscara binária do cérebro do paciente. Além de limitar a análise à região de interesse, a utilização da máscara do cérebro também permite a redução do tempo de processamento.

3.4.1 GLCM

A técnica GLCM, desenvolvida por Haralick (1973), permite a extração de atributos de textura a partir de estatísticas de segunda ordem, condensadas em matrizes de coocorrência de valores de intensidade. Formalmente, para duas dimensões, podemos descrever a técnica GLCM como: dada uma imagem $I(x, y)$ com L níveis de cinza, a matriz de coocorrência $P_{(\Delta x, \Delta y)}$ é uma matriz quadrada $L \times L$, parametrizada por uma distância d e uma direção $\theta = (\Delta x, \Delta y)$, tal que:

$$P_{(\Delta x, \Delta y)}(i, j) = \sum_{p=1}^n \sum_{q=1}^m \begin{cases} 1, & \text{se } I(p, q) = i \text{ e } I(p + d \cdot \Delta x, q + d \cdot \Delta y) = j \\ 0, & \text{caso contrário} \end{cases} \quad (1)$$

em que n e m correspondem ao número de linhas e colunas da imagem I , respectivamente. Cada elemento da matriz $P_{(\Delta x, \Delta y)}$ armazena o número de coocorrências de dois valores de intensidades, i e j , em uma determinada direção θ e separados por uma distância d . Ainda no caso 2D, as orientações são tipicamente definidas como 0° , 45° , 90° e 135° , que

correspondem respectivamente à dupla de deslocamentos $(\Delta x, \Delta y) = (1, 0), (1, 1), (0, 1)$ e $(-1, 1)$. Assim é calculada uma matriz de coocorrência para cada direção, e então cada matriz é normalizada, dividindo cada elemento pela soma de todos os elementos de sua matriz. Por fim, as orientações são combinadas somando-se os elementos das matrizes normalizadas em cada direção e dividindo cada elemento da matriz resultante pela soma de todos os seus elementos, sendo esta representada aqui por $p(i, j)$. Para aplicação das técnicas GLCM e RLM em imagens 3D, mantêm-se a ideia principal descrita nesta seção, porém as direções θ são analogamente escaladas para três dimensões. Sendo os deslocamentos acontecendo nos três eixos como em $\theta = (\Delta x, \Delta y, \Delta z)$ para treze direções.

Com base na matriz GLCM normalizada $p(i, j)$ extraímos 8 atributos de textura em 13 direções. Considerando μ_x, μ_y e σ_x, σ_y , como correspondentes às médias e variâncias, respectivamente, nas linhas e colunas da matriz. Os atributos são calculados conforme descrito a seguir:

O atributo de energia calcula a uniformidade e a organização da textura, quanto maior o valor de energia, maior a uniformidade e organização da textura:

$$F_1 = \sum_{i,j} p(i, j)^2. \quad (2)$$

O atributo de entropia calcula o nível de organização da textura. Quanto mais aleatória for a distribuição de intensidades na imagem, maior será o valor da entropia:

$$F_2 = \begin{cases} \sum_{i,j} p(i, j) \log_2 p(i, j), & \text{se } p(i, j) \neq 0 \\ 0, & \text{se } p(i, j) = 0 \end{cases}. \quad (3)$$

O atributo de correlação calcula a dependência linear entre dos níveis de cinza da GLCM:

$$F_3 = \sum_{i,j} \frac{(i - \mu_x)(j - \mu_y)p(i, j)}{\sigma_x \cdot \sigma_y}. \quad (4)$$

O atributo de momento de diferença inverso (Inverse difference moment - IDM) calcula a homogeneidade da imagem:

$$F_4 = \sum_{i,j} \frac{p(i, j)}{1 + (i - j)^2}. \quad (5)$$

O atributo de inércia, também conhecida como contraste, calcula variações locais dos níveis de cinza da GLCM:

$$F_5 = \sum_{i,j} (i - j)^2 p(i, j). \quad (6)$$

O atributo de sombra de agrupamento (*Cluster Shade*) calcula a assimetria da matriz:

$$F_6 = \sum_{i,j} ((i - \mu_x) + (j - \mu_y))^3 p(i, j). \quad (7)$$

O atributo de proeminência de agrupamento (*Cluster Prominence*) também calcula a assimetria da matriz, quanto maior o valor calculado, menos simétrica a imagem é:

$$F_7 = \sum_{i,j} ((i - \mu_x) + (j - \mu_y))^4 p(i, j). \quad (8)$$

O atributo de correlação de *Haralick* calcula a dependência linear entre os níveis de cinza:

$$F_8 = \frac{\sum_{i,j} (i, j) p(i, j) - (\mu_x \cdot \mu_y)}{\sigma_x \cdot \sigma_y}. \quad (9)$$

3.4.2 RLM

A técnica RLM foi inicialmente proposta por Galloway (1974) e tem sua forma de funcionamento muito similar ao GLCM, considerando os mesmos parâmetros do GLCM de orientação e distância. A técnica RLM caracteriza as texturas de uma imagem pelo número de vezes que dois ou mais voxels de mesmo valor de intensidade são colineares e vizinhos dentro de uma direção θ em uma distância de vizinhança d .

A matriz RLM é calculada de forma similar ao GLCM, tendo uma matriz correspondente por orientação, e cada nível de cinza da imagem representa uma linha da matriz, podendo possuir uma ou mais colunas, sendo que cada coluna representa uma distância d distinta. Considerando uma matriz RLM $P(i, j)$, em que i representa as linhas (número de níveis de cinza) e j as colunas (distâncias da vizinhança de busca), cada elemento da matriz corresponde ao número de vezes que a valor de intensidade de cinza i de uma imagem de I foi colinear numa direção θ numa vizinhança j .

De forma similar ao GLCM, cada matriz $P_{(\Delta x, \Delta y)}(i, j)$ é normalizada gerando a matriz $p(i, j)$, da qual extraímos 10 atributos:

O atributo *Short run emphasis* (SRE) calcula a distribuição de buscas curtas, SRE tem alto valor para texturas mais finas:

$$F_9 = \frac{\sum_{i,j} \frac{p(i,j)}{j^2}}{\sum_{i,j} p(i,j)}. \quad (10)$$

O atributo *Long run emphasis* (LRE) calcula a distribuição de buscas longas, LRE tem alto valor para texturas mais largas:

$$F_{10} = \frac{\sum_{i,j} p(i,j)j^2}{\sum_{i,j} p(i,j)}. \quad (11)$$

O atributo não uniformidade dos níveis de cinza (*Gray level non uniformity* - GLN) calcula a dissimilaridade dos níveis de cinza dentro da textura:

$$F_{11} = \frac{\sum_i (\sum_j p(i, j))^2}{\sum_{i,j} p(i, j)}. \quad (12)$$

O atributo *Run length non uniformity* (RLN) calcula a similaridade dos tamanhos das sequências de níveis de cinza na imagem:

$$F_{12} = \frac{\sum_j (\sum_i p(i, j))^2}{\sum_{i,j} p(i, j)}. \quad (13)$$

O atributo *Low gray level run emphasis* (LGRE) é ortogonal ao SRE:

$$F_{13} = \frac{\sum_{i,j} \frac{p(i,j)}{i^2}}{\sum_{i,j} p(i, j)}. \quad (14)$$

O atributo *High gray level run emphasis* (HGRE) é ortogonal ao LRE:

$$F_{14} = \frac{\sum_{i,j} p(i, j) i^2}{\sum_{i,j} p(i, j)}. \quad (15)$$

O atributo *Short run lowgray level emphasis* (SRLGE) é uma combinação da SRE e LGRE:

$$F_{15} = \frac{\sum_{i,j} \frac{p(i,j)}{i^2 j^2}}{\sum_{i,j} p(i, j)}. \quad (16)$$

O atributo *Short run high gray level emphasis* (SRHGE) é ortogonal a SRLGE e LRHGE:

$$F_{16} = \frac{\sum_{i,j} \frac{p(i,j)}{j^2}}{\sum_{i,j} p(i, j)}. \quad (17)$$

O atributo *Long run low gray level emphasis* (LRLGE) calcula quando a textura é dominada por grandes sequências de baixos níveis de cinza:

$$F_{17} = \frac{\sum_{i,j} \frac{p(i,j) j^2}{i^2}}{\sum_{i,j} p(i, j)}. \quad (18)$$

O atributo *Long run high gray level emphasis* (LRHGE) calcula quando a textura é dominada por grandes sequências de altos níveis de cinza:

$$F_{18} = \frac{\sum_{i,j} p(i, j) i^2 j^2}{\sum_{i,j} p(i, j)}. \quad (19)$$

As implementações das técnicas GLCM e RLM utilizadas nessa pesquisa fazem parte da biblioteca de análise de imagens ITK (Insight ToolKit)⁴. Optamos por utilizar os mesmos parâmetros das implementações para as duas técnicas, que são: distância entre voxels $d = 1$, para obter variações de textura com menor escala nas matrizes, e parâmetros específicos da implementação das técnicas pelo ITK, tal como o número de bins=16 para a quantização de valores de intensidade de cinza na escala de 0 a 255.

⁴ <https://itk.org/>

3.4.3 LBP

A técnica LBP foi inicialmente proposta por Ojala, Pietikäinen e Harwood (1996) como um algoritmo para a descrição de atributos de textura em imagens 2D. O LBP representa a textura a partir de comparações locais, de um pixel/voxel central e sua vizinhança. Para um dado *patch* na imagem, o algoritmo compara o valor do pixel central do *patch* com os pixels de sua vizinhança a fim de definir quais vizinhos possuem valores superiores ou inferiores ao central. Definindo o valor 1 para valores inferiores ao central, e 0 para valores superiores. Este processo gera uma sequência de 1's e 0's formando um número binário. Por fim, este número binário é convertido para base decimal para representar o valor de textura no pixel correspondente na imagem de saída.

Como as imagens utilizadas nesta pesquisa são tridimensionais, recorreremos a uma abordagem do LBP que atua com imagens 3D. Utilizamos a extensão para 3D proposta por Banerjee et al. (2012), que define a vizinhança de comparação por esferas harmônicas com intuito de obter um método invariante à rotação. O método LBP 3D requer o ajuste de três parâmetros que são: (a) o número de níveis de esferas harmônicas (n_{shl}), (b) o raio (r_n) em que os vizinhos devem ser amostrados e (c) o número de subdivisões (n_i) para aplicar na icosfera. Ao alterar esses parâmetros, podemos obter imagens de textura diferentes.

Com base na análise experimental usando um subconjunto de dados de cinquenta imagens RM (25 CN e 25 AD), conforme mencionado na Seção 3.2, descobrimos que os valores dos parâmetros $n_{shl} = 3$, $r_n = 2$, e $n_i = 1$ fornecem o melhor compromisso entre resultados de alta acurácia e baixo custo computacional. Além das imagens de textura resultantes com os parâmetros descritos acima, o método fornece uma imagem de *kurtosis* que também foi incorporada em nossa análise. Portanto, o método LBP fornece um total de quatro atributos de textura, definidos aqui como F_{19} para o primeiro nível de esferas harmônicas, F_{20} para o segundo nível, F_{21} para o terceiro e F_{22} para a *kurtosis*.

3.5 Extração de características de textura

A aplicação dos três filtros de textura em cada imagem dos pacientes produziu 22 imagens filtradas, sendo elas 8 GLCM, 10 RLM e 4 LBP. Neste trabalho, extraímos e representamos os atributos de textura utilizando duas abordagens. A primeira abordagem representa cada região segmentada pela distribuição de probabilidade obtida da normalização dos histogramas de textura de cada região. Neste estudo, cada histograma teve as larguras dos bins fixadas, resultando num número fixo de bins. Essas distribuições foram utilizadas para calcular o peso das arestas entre vértices dos grafos. A segunda abordagem extrai os quatro primeiros momentos estatísticos dos valores de textura de cada região segmentada e concatena esses valores com o volume da correspondente região. Assim,

após a etapa de filtragem de texturas, representamos cada paciente s por uma matriz de características M , que é definida como:

$$M^s = \begin{bmatrix} f_{0,0} & f_{0,1} & \cdots & f_{0,n_f-1} \\ f_{1,0} & f_{1,1} & \cdots & f_{1,n_f-1} \\ \vdots & \vdots & \ddots & \vdots \\ f_{n_r-1,0} & f_{n_r-1,1} & \cdots & f_{n_r-1,n_f-1} \end{bmatrix}, \quad (20)$$

sendo $n_r = 92$ a quantidade de regiões segmentadas e $n_a = 22$ a quantidade de imagens de textura (atributos de textura de resposta aos filtros). Cada linha da matriz representa uma região segmentada e as colunas indicam correspondem às imagens filtradas de textura.

A matriz M^s é como os pacientes são representados desta etapa em diante, e as modificações e abordagens de extração de características são executados em cada matriz. Na Figura 4 apresentamos uma representação visual das etapas usadas para obter a matriz de atributos M de um paciente, conforme apresentado na Figura 2.

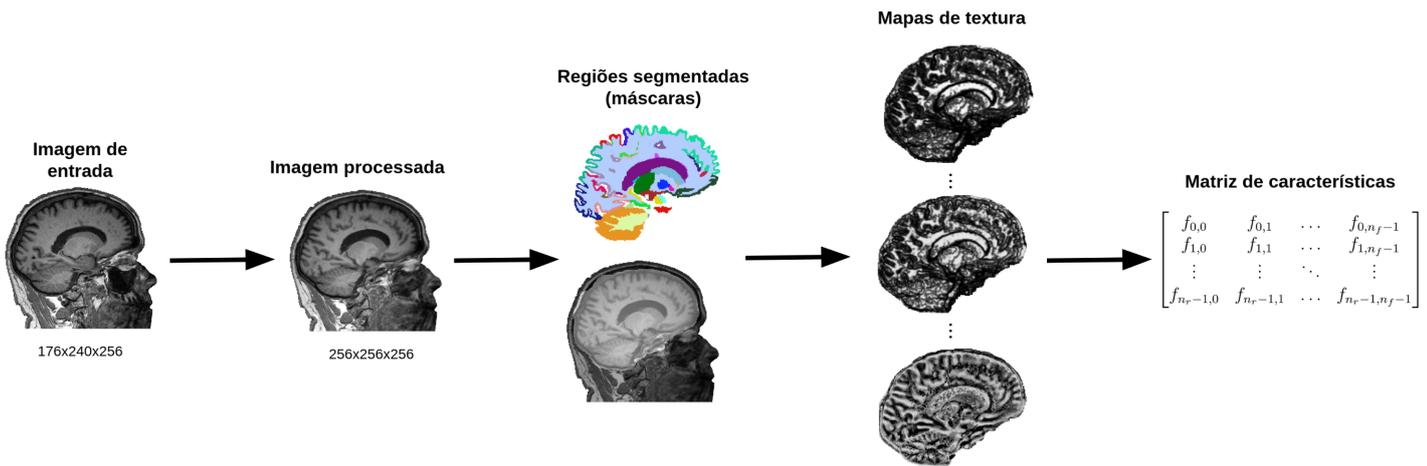


Figura 4 – A representação visual das primeiras etapas da Figura 2 para um paciente.

Usando o FreeSurfer, cada imagem da ADNI é registrada a um novo espaço e uma máscara binária do cérebro é extraída. Uma vez que a imagem do paciente é registrada ao espaço estereotáxico do FreeSurfer, as máscaras das regiões cerebrais segmentadas com dois atlas ficam disponíveis para uso. A seguir, os filtros de textura GLCM, RLM e LBP são então aplicados às imagens pré-processadas para extrair os atributos de textura. Finalmente, cada elemento da matriz de atributos M^s é obtido observando-se às regiões segmentadas e os atributos de textura gerados.

3.5.1 Normalização

Os atributos de textura foram normalizados entre todos os pacientes utilizando a técnica min-max. A normalização min-max efetua uma transformação linear nos dados para que eles fiquem espalhados dentro de um novo intervalo de dados, possibilitando a

comparação de valores de textura entre os pacientes. A normalização min-max utilizada seguiu a definição apresentada por Han, Kamber e Pei (2012) em que, a partir do valor máximo e mínimo entre todos os pacientes, a normalização é calculada como:

$$v_{norm}^{(i)} = \frac{v^{(i)} - \min_A}{\max_A - \min_A} (\text{novo_max}_A - \text{novo_min}_A) + \text{novo_min}_A, \quad (21)$$

sendo A o conjunto de todos os valores de intensidade de um atributo extraído para todos os pacientes e $v^{(i)}$ e $v_{norm}^{(i)}$ os valores de um atributo $i \in A$ antes e após a normalização.

Partindo da matriz M^s da equação 20, pode-se analisar a variação de intensidade dos atributos entre todos os pacientes. Considerando a possível variação de intensidades de textura, e dado que a filtragem é realizada individualmente para cada paciente, a normalização é necessária para fins de comparação e posterior classificação das imagens. Logo, a normalização min-max de todos os valores de textura é realizada para o intervalo $[0.1, 1]$. Os valores de mínimo e máximo são extraídos entre todos os pacientes, individualmente para cada atributo de textura a partir de todas as matrizes M^s (Eq. 20). A normalização min-max também é necessária para manter a consistência dos valores de atributos entre os pacientes, visto que a construção dos grafos e cálculo da métrica de similaridade dependem da distância entre um mesmo atributo de textura.

3.5.2 Histogramas e distribuições de texturas

Visto que o número de valores de textura variam conforme a quantidade de voxels da ROI, nesta pesquisa representamos cada região segmentada e filtrada por uma distribuição de probabilidade obtida da normalização do histograma de textura. Podemos entender a representação via histograma como um mapeamento de dados de alta dimensão para um espaço de dimensionalidade menor (BEHESHTI et al., 2017). Porém, não há método universal para a escolha do número de *bins* e suas larguras, ficando a cargo do pesquisador experimentar e decidir quais parâmetros melhor representa os dados da análise (DIECKMEYER et al., 2021).

Como cada região segmentada possui um número de voxels diferente, faz-se necessário o uso de um método para calcular o número de *bins* ótimo para cada atributo de textura. Existem diversos algoritmos para este cálculo, por exemplo Hett et al. (2018) utilizou a regra de Sturge (STURGES, 1926), porém no seu caso o valor de cada voxel tinha uma variação restrita entre -1 a 1, isto é, não havia a possibilidade de *outliers*. Entretanto com o uso de filtros de textura, cada atributo calculado possui uma grande variação de valores, suscetível a *outliers* e com intervalos diferentes para cada imagem, o que resulta em histogramas potencialmente distintos. Para solucionar este problema, utilizamos a técnica Freedman-Diaconis (FREEDMAN; DIACONIS, 1981), que é uma técnica já utilizada por outros trabalhos publicados na literatura para atributos de textura do GLCM (DIECKMEYER et al., 2021; NOORTMAN et al., 2020). Seu cálculo leva em consideração a

amplitude interquartil e a quantidade de dados, resultando em robustez à *outliers*. O comprimento dos *bins* h pela regra de Freedman-Diaconis é calculado conforme a equação a seguir:

$$h = \frac{2 \cdot IQR}{\sqrt[3]{N}}, \quad (22)$$

e

$$\#bins = \left\lfloor \frac{\max(I_{ROI}) - \min(I_{ROI})}{h} \right\rfloor, \quad (23)$$

sendo N o número de voxels de uma região segmentada (I_{ROI}) e IQR a sua amplitude interquartil.

Para determinar a quantidade de bins para cada atributo de textura, utilizamos os 50 pacientes (25 DA e 25 CN) que foram selecionadas exclusivamente para a estimação de parâmetros (Seção 3.2) como o número de bins por atributo de textura. Usando as matrizes de características dos 50 pacientes, calculamos o histograma para cada região e cada atributo de textura utilizando a regra de Freedman-Diaconis. Esse procedimento representa operações nas colunas da matriz de características M^s . Finalmente, obtemos o valor médio de bins de todas as regiões para cada atributo de textura dentre os 50 pacientes, o que resulta em 22 números de bins, um para cada atributo de textura. Essa etapa gera uma nova matriz de características para cada paciente, a qual é composta pela distribuição de probabilidade dos valores de textura de cada região (linha) e para cada atributo de textura (coluna). As distribuições são todas estimadas a partir dos seus respectivos histogramas com o número de bins fixo. Assim, os elementos da matriz M^s (Eq. 20) podem ser definidos como

$$f_{i,j} = [distribuição-de-probabilidade_{i,j}], \quad (24)$$

em que i e j correspondem, respectivamente, aos índices da região segmentada do cérebro e do atributo de textura. Após a definição desta nova matriz, normalizamos com a técnica de normalização min-max (Seção 3.5.1) todas as distribuições por região dentre todos os pacientes. Essa normalização representa uma operação nas linhas das matrizes.

Ao fixar o número de bins dos histogramas por atributo, regiões muito pequenas ou muito grandes podem influenciar negativamente na média de bins, atuando como outliers para a distribuição de número de voxels das regiões. Conforme descrito na seção 3.3.2 removemos da análise, as regiões com número médio de voxels entre pacientes inferior à menor região segmentada relevante à doença de Alzheimer, o córtex entorhinal (EN-KIRCH et al., 2018). Além disso, removemos também as regiões com número médio de voxels superior a 25.000, a qual corresponde a toda a região de WM.

3.5.3 Volume da região e momentos estatísticos

Como uma segunda forma de representação dos atributos de textura por região, utilizamos o volume das regiões segmentadas e os quatro primeiros momentos estatísticos das distribuições calculadas na subseção anterior. A doença de Alzheimer é conhecida por causar um processo de atrofia em diversas regiões cerebrais, e esta atrofia pode ser medida a partir do volume das regiões, quando comparadas aos pacientes do grupo controle (CN). Os valores de volume das regiões segmentadas foram obtidos a partir das regiões segmentadas pelo FreeSurfer (FRENZEL et al., 2020; WEE et al., 2019).

Então, como uma estratégia auxiliar às distribuições de texturas, criamos um vetor que combina o volume da região com os quatro primeiros momentos estatísticos dos atributos de textura da região. Assim, uma nova matriz M^s é construída e seus elementos são definidos como:

$$f_{i,j} = [\text{volume}_i, \text{média}_{i,j}, \text{variância}_{i,j}, \text{skewness}_{i,j}, \text{kurtosis}_{i,j}], \quad (25)$$

sendo que i e j correspondem, respectivamente, aos índices da região segmentada no cérebro e do atributo de textura analisado.

3.6 Construção dos grafos cerebrais

Os grafos utilizados neste estudo são definidos como: $G = (V, E, a)$, sendo V o conjunto de vértices, com $|V| = n$ representando o número de vértices, E o conjunto de arestas (uma aresta é a conexão entre dois vértices) ponderadas tal que $(u, v) \in A$ para $u, v \in V$, e uma função de atribuição aos vértices $a : V \rightarrow \mathbb{R}^D$, onde D é a dimensão de um vetor de atributos contínuos. Os grafos utilizados nesta pesquisa são do tipo não-direcionados e sem laços (aresta ligando com o próprio vértice), ou seja, $(1, 2) = (2, 1)$.

A Figura 5 ilustra um exemplo de grafo com três vértices, 1, 2 e 3. Cada vértice possui um vetor de atributos de três dimensões e com duas arestas $(1, 2)$ e $(2, 3)$ com pesos 0.42 e 0.27 respectivamente.

Um grafo pode ser representado por uma matriz de adjacência, que é uma matriz quadrada $n \times n$, denotada por $A(G) = [a_{ij}]$ e com seus elementos definidos pela seguinte regra,

$$a_{ij} = \begin{cases} d(v_i, v_j), & \text{se existir uma aresta entre os vértices } v_i \text{ e } v_j \\ 0, & \text{caso contrário} \end{cases}, \quad (26)$$

sendo $d(v_i, v_j)$ a distância entre dois vértices v_i e v_j , que é calculada a partir dos atributos dos respectivos vértices. Para o grafo da Figura 5, a sua matriz de adjacência é dada por

$$A(G) = \begin{bmatrix} 0 & 0.42 & 0 \\ 0.42 & 0 & 0.27 \\ 0 & 0.27 & 0 \end{bmatrix}.$$

Com o objetivo de modelar o cérebro a partir das regiões segmentadas, iniciamos com grafos completos, isto é, todo vértice é adjacente a todos os demais vértices. Considerando um conjunto de vértices V , tal que $n_v = |V| =$ número de ROI's, podemos então calcular a quantidade de arestas do grafo pela seguinte combinação simples de dois elementos,

$$|E| = C(n_v, 2) = \frac{n_v!}{2!(n_v - 2)!} = \frac{n_v(n_v - 1)(n_v - 2)!}{2!(n_v - 2)!} = \frac{n_v(n_v - 1)}{2}, \quad (27)$$

e, como nesta pesquisa estamos utilizando $n_v = 92$ (número de regiões segmentadas), então o número de arestas é $|E| = 4186$.

Conforme definido na seção 3.4, para cada imagem extraímos 22 mapas de textura de todas as 92 regiões do cérebro. Para cada mapa de textura, construímos dois tipos de grafos. No primeiro tipo, os vetores de atributos dos vértices do grafo, representando as regiões segmentadas da imagem, foram definidos pelas distribuições de textura, enquanto no segundo tipo os atributos dos vértices foram definidos pelos vetores compostos pelos quatro primeiros momentos estatísticos dos níveis de cinza e pelo volume de cada região segmentada (Seção 3.5.3). Os detalhes da construção do gráfico são fornecidos nas seções subsequentes.

3.6.1 Atributos dos vértices

Conforme mencionado anteriormente, desenvolvemos duas abordagens para representar os atributos de cada região cerebral. O objetivo foi criar um grafo cujos vértices possuam atributos que represente muito bem suas respectivas regiões.

A primeira forma que utilizamos para isso foi inserir diretamente a distribuição de probabilidade dos valores de textura da região como atributos do vértice, e a segunda forma, foi o vetor contendo o volume da região e os quatro primeiros momentos estatísticos da mesma. Podemos sumarizar os dois tipos de grafos criados pela perspectiva do tipo de atributo de vértices, como:

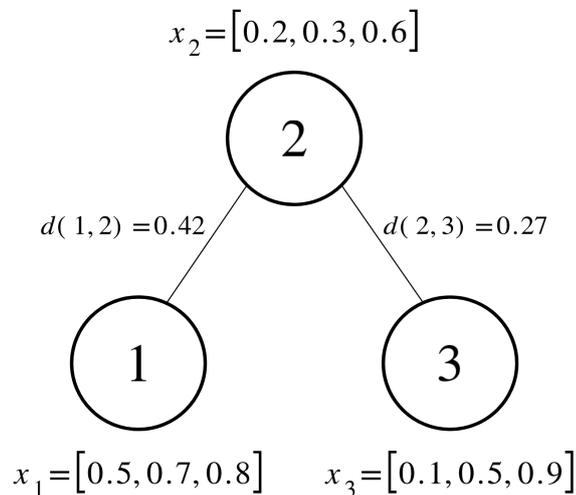


Figura 5 – Exemplo de grafo G com atributos nos vértices.

1. A distribuição dos valores do atributo de textura (Equação 24).
2. O vetor de atributos composto pelo volume da região e os momentos estatísticos (Equação 25).

3.6.2 Definição das arestas e seus pesos

Considerando a primeira abordagem de extração de atributos (Seção 3.5.2), que representa os valores de textura de cada região filtrada como uma distribuição de probabilidades, podemos então calcular a dissimilaridade entre regiões por meio da distância entre suas distribuições. Com o objetivo de obter um valor de dissimilaridade consistente entre duas regiões, foi preciso escolher métricas capazes de calcular de forma simétrica a distância entre as distribuições de probabilidade de textura das regiões, dado que nosso grafo é não-direcionado. Logo, selecionamos três distâncias conhecidas e bem estabelecidas na literatura que são: a distância de Wasserstein, a de Kullback-Leibler e a a distância de Hellinger.

3.6.2.1 Wasserstein

A distância de Wasserstein (D_{ws}), também conhecida como *Earth Mover Distance* (RUBNER; TOMASI; GUIBAS, 2000) pode ser entendida, intuitivamente, como o custo de transportar terra de uma montanha para outra, considerando que cada montanha corresponde a uma distribuição. Essa métrica foi aplicada no trabalho de Hett et al. (2018), conforme descrito no Cap. 2. Nesse estudo, utilizamos a formulação equivalente proposta por Ramdas, Trillos e Cuturi (2017), em que a distância de Wasserstein para duas distribuições de mesma área u e v é definida como:

$$D_{ws}(u, v) = \sum |cdf(u) - cdf(v)|, \quad (28)$$

sendo $cdf(\cdot)$ é a função de distribuição acumulada

3.6.2.2 Kullback-Leibler

A distância de Kullback-Leibler (D_{kl}) (KULLBACK; LEIBLER, 1951), também conhecida como entropia relativa da teoria da informação, uma interpretação é pela a quantidade de informação perdida ao aproximar uma distribuição de probabilidade para outra. Em que a D_{kl} se aproxima de zero quando as duas distribuições possuem informações similares. Embora esta distância não seja simétrica, neste estudo a propriedade de simetria foi obtida utilizando a distância média entre $D_{kl}(u, v)$ e $D_{kl}(v, u)$, como

$$D_{sym-kl}(u, v) = \frac{1}{2} [D_{kl}(u||v) + D_{kl}(v||u)], \quad (29)$$

em que

$$D_{kl}(p||q) = \sum_{x \in \mathcal{X}} p(x) \ln \frac{p(x)}{q(x)}, \quad (30)$$

sendo x uma instância do espaço de probabilidades \mathcal{X} .

3.6.2.3 Hellinger

A distância de Hellinger (D_{hg}), inicialmente proposta por Hellinger (1909), é semelhante à distância *Total Variation*, exceto que se assemelha à norma L2. Em sua forma discreta, a distância é calculada como

$$D_{hg}(u, v) = \frac{1}{\sqrt{2}} \left\| \sqrt{u} - \sqrt{v} \right\|_2. \quad (31)$$

3.6.3 Limiarização

Um dos principais desafios de trabalhar com grafos construídos a partir de imagens de RM, é como construir grafos discriminativos para uso em uma tarefa de classificação. Na literatura, alguns trabalhos usam o coeficiente de correlação de Pearson (LIU et al., 2016; LIU et al., 2017; JIE et al., 2016) para definir um valor de limiar mínimo com base no p-valor, enquanto outros trabalhos usam a frequência de *minimum spanning tree* entre os vértices para encontrar o sub-grafo mais frequente (CUI et al., 2018). Uma abordagem comum em vários trabalhos da literatura ao trabalhar com grafos cerebrais é iniciar com um gráfico completo, que é o caso em que todos os vértices estão conectados uns aos outros, e então aplicar um método de remoção de arestas para obter uma representação de um grafo mais discriminativo. Neste estudo desenvolvemos uma abordagem baseada em limiares para a remoção de arestas dos grafos que gera um valor de limiar para cada mapa de textura.

Neste estudo iremos considerar dois grupos de grafos completos, G_{CN} e G_{DA} , com o mesmo conjunto de vértices V , tal que $G_{CN}^{(i)} \in G_{CN}$ e $G_{DA}^{(i)} \in G_{DA}$ para $i = 1, \dots, n_s$, e i o enésimo grafo de cada conjunto. A quantidade de grafos é $n_s = 25$ o que corresponde ao número de imagens selecionadas para a estimação de parâmetros por grupo de diagnóstico (DA e CN), conforme mencionado na seção 3.2. A partir desses dois grupos de grafos podemos calcular um peso médio e um desvio padrão médio entre todos os grafos do grupo de estimação de parâmetros. Lembrando que o peso das arestas é calculado pela distância entre as distribuições de probabilidade das respectivas regiões/vértices. Portanto, quanto maior a distância, maior será a diferença entre as texturas das regiões dos vértices.

Considere $\mathbf{d}_f \in \mathbb{R}^{n_s}$ como um vetor que possui a distância média entre pares de grafos dos conjuntos G_{CN} e G_{DA} (comparações de pares) como seus elementos, e considere f como sendo um índice dos atributos definidos na seção 3.4, tal que $f \in [1, \dots, 22]$, então

os elementos de \mathbf{d}_f são calculados como

$$d(G_{CN}^{(i)}, G_{DA}^{(i)})_{f,i} = \frac{\sum_{e \in E} |w(G_{CN}^{(i)}, e) - w(G_{DA}^{(i)}, e)|}{|E|}, \quad (32)$$

em que $e = (u, v) \in E$ é uma aresta formada pelos vértices $u, v \in V$, $|E|$ é o tamanho do conjunto de arestas, e $w : G, e \rightarrow \mathbb{R}$ é uma função de peso que recebe um grafo G e uma aresta e e tem como saída o peso da aresta para o grafo, que é calculado pela distância entre as distribuições de textura dos vértices. Nesse caso, o peso das arestas pode ser calculado pelas três distâncias definidas na seção 3.6.2.

Para determinar a representação dos grafos mais discriminativos para a classificação das imagens, definimos um limiar T_f

$$T_f(n) = \mu_f + n \cdot \sigma_f, \quad (33)$$

para cada mapa de textura f que é o valor médio das distâncias entre dois grupos de diagnóstico armazenados no vetor \mathbf{d}_f , em que n é um fator multiplicativo, e μ_f e σ_f correspondem a média e o desvio padrão, respectivamente, para um mapa de textura f qualquer, e são calculados como

$$\mu_f = \frac{\sum_{i=1}^{n_s} \mathbf{d}_{f,i}}{n_s}, \quad (34)$$

e

$$\sigma_f = \sqrt{\frac{\sum_{i=1}^{n_s} |\mathbf{d}_{f,i} - \mu_f|^2}{n_s}}. \quad (35)$$

Neste caso, quanto maior o valor de n , mais arestas serão removidas e, caso um vértice tenha todas suas arestas removidas, ele será removido. Os melhores valores de n são determinados experimentalmente e apresentados no Capítulo 4 para os valores $n = 0.5, 1, 2, 3, 4, 5$. A partir deste método, foi possível extrair um valor de limiar para cada mapa de textura e assim construir grafos com uma melhor separação entre classes.

3.7 Graph Kernel

Como o grafo é uma estrutura complexa sem representação direta que pode ser útil para classificação, o *Graph Kernel* (GK) é uma das principais abordagens para resolver o desafio que é calcular a similaridade entre dois ou mais grafos complexos (RICHIARDI et al., 2013b; KRIEGE; JOHANSSON; MORRIS, 2020b). Os métodos de GK podem ser adaptados a diferentes tipos de problema a partir da perspectiva das propriedades dos grafos envolvidos. O GK atua como um facilitador para a representação vetorial de grafos e para o cálculo da similaridade entre pares de grafos.

Métodos de *kernel* são particularmente utilizados em algoritmos de aprendizado de máquina, como o classificador SVM, para aprender a partir de medidas de similaridade entre os dados (SCHÖLKOPF et al., 2002). Considerando um conjunto de dados $\chi \in \mathbb{R}^d$

e definindo $k : \chi \times \chi \rightarrow \mathbb{R}$ uma função, então a função k é um *kernel* se existir no espaço de Hilbert \mathcal{H}_k um mapeamento de atributos $\phi : \chi \rightarrow \mathcal{H}_k$ tal que $k(x, y) = \langle \phi(x), \phi(y) \rangle$ para $x, y \in \chi$, em que $\langle \cdot, \cdot \rangle$ é o produto interno em \mathcal{H}_k . O mapeamento ϕ existe, se e somente se, k for uma função positiva semidefinida. Para exemplificar, podemos utilizar o caso em que $\chi = \mathbb{R}^d$ e $\phi(\mathbf{x}) = \mathbf{x}$, resultando no produto escalar $k(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{y}$. (SCHÖLKOPF et al., 2002)

O *kernel* também possui uma propriedade que permite acessar os dados diretamente pela matriz de Gram K , isto é, não é necessário uma representação explícita dos dados de χ como um vetor $\phi(\mathbf{x})$, sendo então uma ótima forma para se trabalhar com grafos. A matriz K é definida a partir de um conjunto de dados $x_1, \dots, x_m \in \chi$, com elementos K_{ij} , para $i, j \in 0, \dots, m$, correspondendo ao resultado do *kernel* entre pares de dados de χ , ou seja, $K_{ij} = k(x_i, x_j)$. Se a matriz K for positiva e semidefinida para todos os possíveis pares de dados, então k é um *kernel* (SCHÖLKOPF; SMOLA; MÜLLER, 1997) e pode ser representado como

$$K(x_1, \dots, x_n) = \begin{bmatrix} k(x_1, x_1) & k(x_1, x_2) & \dots & k(x_1, x_n) \\ k(x_2, x_1) & k(x_2, x_2) & \dots & k(x_2, x_n) \\ \vdots & \vdots & \ddots & \vdots \\ k(x_n, x_1) & k(x_n, x_2) & \dots & k(x_n, x_n) \end{bmatrix}. \quad (36)$$

A partir da matriz de Gram da equação 36, e considerando a propriedade descrita no parágrafo anterior, um *Graph Kernel* é uma função $k : \mathbb{G} \times \mathbb{G} \rightarrow \mathbb{R}$ em que \mathbb{G} indica um conjunto de grafos.

Nessa pesquisa utilizamos dois *graph kernels*, o *Propagation Kernel* (P2K) (NEUMANN et al., 2016) e o *Hash Graph Kernel* (HGK) (MORRIS et al., 2016), sendo que ambos os kernels podem ser utilizados com grafos não-direcionados, com seus vértices representados por vetores d -dimensionais. Essa abordagem de *kernel* permite que o classificador SVM use diretamente em sua formulação a matriz de Gram K gerada pelo *Graph Kernel*.

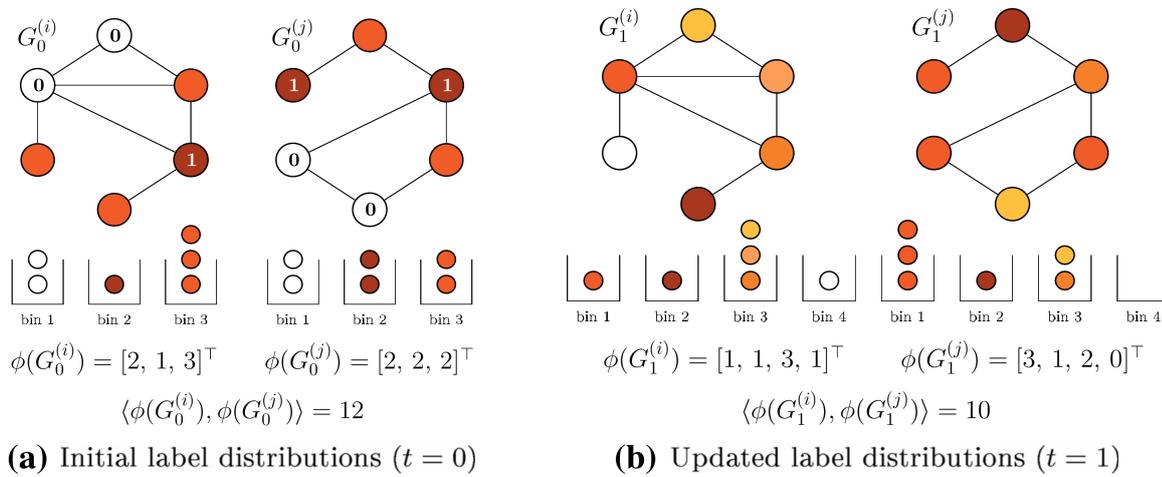
3.7.1 Propagation Kernel

O *kernel* P2K foi proposto como um *graph kernel* multipropósito, visando suprir as limitações de outros *kernels* existentes na literatura, uma vez que possui a capacidade de trabalhar com grafos que possuem vetores de d dimensões como atributo dos vértices (KRIEGE; JOHANSSON; MORRIS, 2020b). A ideia geral de seu funcionamento consiste em mapear a distribuição dos vetores de atributos de cada vértice de um grafo para um novo espaço, a fim de possibilitar o cálculo da similaridade de distribuições de atributos entre grafos neste novo espaço. Em termos gerais, um *kernel* P2K entre dois grafos pode ser definido como a comparação entre todos os pares de vértices nos dois grafos, com u e

v sendo vértices dos grafos $G^{(i)}$ e $G^{(j)}$ respectivamente, ou seja

$$K(G^{(i)}, G^{(j)}) = \sum_{v \in G^{(i)}} \sum_{u \in G^{(j)}} k(u, v). \quad (37)$$

Para obter a representação dos atributos dos vértices do grafo, o *kernel* P2K utiliza o algoritmo *locality sensitive hashing* (LSH), proposto por Datar et al. (2004). O algoritmo LSH é capaz de quantizar vetores d -dimensionais a partir de sua proximidade, neste caso a quantização segue um esquema de *binning* pela probabilidade de dois vetores serem próximos, com base na distância entre eles.



Fonte: Neumann et al. (2016)

Figura 6 – Exemplo de funcionamento do *kernel* P2K para dois grafos com atributos de uma dimensão.

A Figura 6 ilustra o funcionamento da quantização e o cálculo de similaridade entre grafos por meio do algoritmo LSH em duas iterações do algoritmo, sendo esse um exemplo em que cada vértice possui um valor numérico como atributo. Nesse caso, o algoritmo calcula a proximidade entre os atributos dos vértices pelo produto escalar do atributo x do vértice com um vetor calculado pela distribuição de Cauchy (CAUCHY, 1853) sobre todos os atributos do grafo e de um deslocamento b calculado a partir de w , que representa a largura dos bins. Por fim, h indica em qual *hash* (bin) o atributo de um vértice será quantizado, ou seja,

$$h(\mathbf{x}; w, \mathbf{v}, b) = \left\lfloor \frac{\mathbf{v}^T \mathbf{x} + b}{w} \right\rfloor. \quad (38)$$

Como último passo do algoritmo, calcula-se o produto escalar entre cada par de vetores $\phi(G^{(i)})$ e $\phi(G^{(j)})$ para construir a matriz de Gram.

3.7.2 Hash Graph Kernel

A principal ideia do HGK é efetuar o *hash* iterativo de atributos contínuos dos vértices para rótulos discretos, o HGK também permite o uso de *kernels* base (implementação

implícita) diferentes para lidar com os atributos derivados do *hash*. Como *kernels* de base, o HGK utiliza variantes dos *Graph Kernels Weisfeiler-Lehman* (WL) *subtree kernel* (SHERVASHIDZE et al., 2011) e do *shortest-path kernel* (SP) (BORGWARDT; KRIEGEL, 2005) para permitir o uso de vetores de atributos nos vértices dos grafos. O HGK entre dois grafos G_1 e G_2 é definido como,

$$HGK(G_1, G_2) = \frac{1}{|\mathfrak{H}|} \sum_{i=1}^{|\mathfrak{H}|} k(\mathfrak{h}_i(G_1), \mathfrak{h}_i(G_2)), \quad (39)$$

em que k é o *kernel* base, tal como o WL e o SP, e $\mathfrak{H} = \{\mathfrak{h}_1, \mathfrak{h}_2, \dots\}$ é uma família de funções *hash* produzidas pelo LSH, sendo cada elemento $\mathfrak{h}_i \in \mathfrak{H}$ uma função, tal que $\mathfrak{h}_i : \mathbb{R}^d \rightarrow \mathbb{N}$. Assim, o HGK aplica iterativamente uma função de *hash* do conjunto \mathfrak{H} no vetor de atributos dos vértices até obter uma baixa variância. A principal ideia do *graph kernel* SP é decompor os grafos em seus *shortest-paths* e comparar os pares desses caminhos de acordo com seu comprimento e atributo dos vértices. O *kernel* SP que é também usado como *kernel* base k é definido como,

$$k_{SP}^{k_A, k_d}(G_1, G_2) = \sum_{(u,v) \in E_1, u \neq v} \sum_{(w,z) \in E_2, w \neq z} k((u,v), (w,z)), \quad (40)$$

sendo

$$k((u,v), (w,z)) = k_A(a(u), a(w)) \cdot k_A(a(v), a(z)) \cdot k_d(d_{uv}, d_{wz}), \quad (41)$$

com d_{uv} sendo o tamanho do caminho mais curto entre u e v para $(u,v) \in E$, k_A sendo um *kernel* para comparação de atributos dos vértices e k_d um *kernel* para comparar as distâncias de menor caminho (*shortest-path*). Nesta pesquisa, também foi utilizada uma adaptação do algoritmo usado pelo SP para permitir que a matriz de adjacência do grafo possua como elementos os pesos das arestas. Neste caso, nomeamos esta modificação como *weighted-SP* (wSP).

A ideia do *graph kernel* WL é substituir o *label* (o atributo do vértice) de cada vértice por um *multiset label*, que consiste do *label* original do vértice e do conjunto de *labels* de seus vizinhos. O conjunto resultante é então compactado em um novo *label* curto. Esse procedimento é repetido para as h repetições. Considerando k_δ como o *kernel* de Dirac e l uma função *label*, então o *kernel* WL pode ser definido como,

$$k_{WL}^h(G_1, G_2) = \sum_{i=1}^h \sum_{e \in E_1, e' \in E_2} k(e, e'), \quad (42)$$

onde,

$$k_i(e, e') = \begin{cases} k_\delta(l(e), l(e')) & i = 0 \\ k_{i-1}(e, e') \cdot f(e, e') & i > 0 \wedge \mathcal{M}_i(e, e') \neq \emptyset, \\ 0 & i > 0 \wedge \mathcal{M}_i(e, e') = \emptyset \end{cases} \quad (43)$$

e

$$f(e, e') = |\mathcal{M}_i(e, e')|^{-1} \sum_{R \in \mathcal{M}_i(e, e')} \prod_{(w, w') \in R} k_{i-1}(w, w'), \quad (44)$$

com $\mathcal{M}_i(e, e')$ sendo o conjunto de bijeções $b : E_1 \rightarrow E_2$ entre a vizinhança de e e e' .

3.8 Esquema de classificação

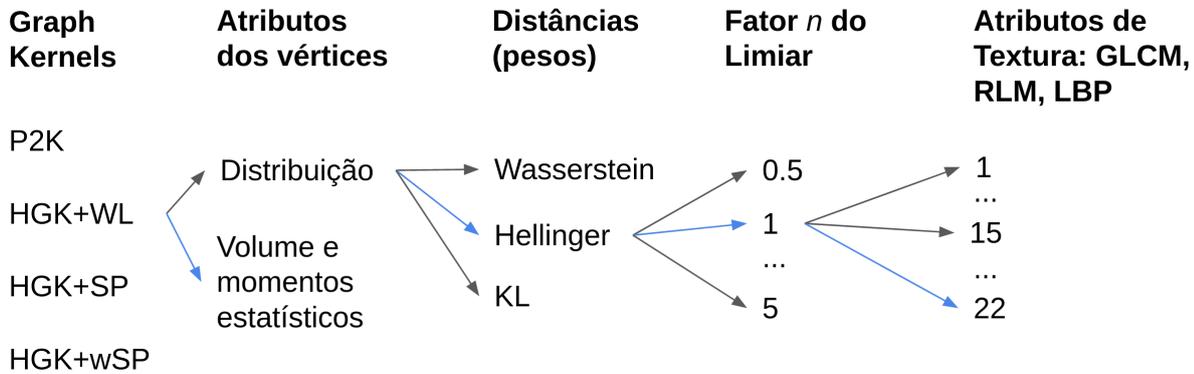


Figura 7 – Esquema de classificação utilizado neste estudo.

Considerando que neste ponto temos quatro *graph kernels*, dois tipos de vetores de atributos para os vértices dos grafos, três distâncias entre distribuições dos vértices, seis valores de n do limiar (Equação 33) e 22 atributos de textura extraídos com GLCM, RLM e LBP, então foi preciso de uma abordagem de classificação para encontrar a melhor combinação desses pontos e assim obter o melhor resultado. Dessa maneira, uma matriz de Gram foi gerada para cada combinação possível desses elementos. Todas as possibilidades de combinação estão representadas na Figura 7, e todas essas combinações foram classificadas para os casos $CN \times AD$, $CN \times MCI$ e $MCI \times AD$.

Em resumo, o classificador SVM recebe a matriz de Gram gerada pelo *graph kernel* para efetuar sua classificação com a matriz de *kernel* pré-computada (mais detalhes podem ser encontrados no apêndice A.2). O desempenho do método foi avaliado aplicando o SVM com validação cruzada de 10-folds usando a acurácia como métrica de sucesso. Para obter resultados com um menor enviesamento, a classificação de cada matriz de Gram foi realizada 10 vezes, assim os resultados finais são extraídos a partir da média aritmética e desvio padrão do resultado das 10 execuções da validação cruzada de 10-folds. E para uma melhor avaliação dos resultados dos experimentos de classificação, também foi computado a sensibilidade, especificidade e AUC.

A Figura 7 ilustra o esquema de classificação utilizado neste trabalho. Para cada *graph kernel*, construímos um grafo para cada abordagem de atributos de vértices e, para cada tipo de atributo de vértice, o peso das arestas são calculados usando as três distâncias (Seção 3.6.2). Assim, para cada distância é aplicado um valor diferente de limiar a partir da variação do parâmetro n para a remoção das arestas dos grafos e esses grafos são montados a partir de cada atributo de textura gerado pelos filtros GLCM, RLM e LBP. Para cada caminho indicado na figura, temos uma matriz de Gram gerada que é então utilizada pelo classificador SVM. Na figura foi destacado um exemplo de caminho que

corresponde ao *graph kernel* HGK+WL, utilizando como atributos Vol+Stats, para a distância de Hellinger, com o parâmetro $n = 1$ e para o atributo de textura F_{22} .

3.9 Métricas de avaliação dos resultados

Para a classificação das imagens de RM entre os grupos de diagnóstico CN, MCI e DA, foram utilizadas as seguintes métricas: acurácia (Acc), sensibilidade (Sens), especificidade (Espec) e a AUC. Essas métricas são calculadas a partir dos verdadeiros positivos (VP), dos falsos positivos (FP), dos verdadeiros negativos (VN) e dos falsos negativos (FN) da matriz de contingência. A partir desses quatro valores, obtidos pelo teste de classificação, a acurácia é definida como

$$\text{Acc} = \frac{\text{VP} + \text{VN}}{\text{VP} + \text{FP} + \text{VN} + \text{FN}}, \quad (45)$$

a qual mede a quantidade total de acertos do classificador.

A sensibilidade calcula a porcentagem de acertos na classe positiva e a especificidade calcula a porcentagem de acertos na classe negativa, já a precisão mede, dentre todas as classificações de classe positiva que o classificador fez, quantas estão corretas. As métricas descritas são definidas como

$$\text{Sens} = \frac{\text{VP}}{\text{VP} + \text{FN}}, \quad (46)$$

$$\text{Espec} = \frac{\text{VN}}{\text{VN} + \text{FP}}, \quad (47)$$

$$\text{Precisão} = \frac{\text{VP}}{\text{VP} + \text{FP}}. \quad (48)$$

3.10 Considerações gerais da metodologia

Este capítulo apresentou a metodologia e as técnicas empregadas neste trabalho de pesquisa, descrevendo cada um dos principais métodos e apresentando uma breve justificativa de suas escolhas para este trabalho. Podemos também ressaltar que existem algumas limitações em relação à metodologia empregada nesta pesquisa, como:

- a) A dependência da precisão da técnica de segmentação utilizada, visto que a extração dos atributos de textura serão realizados individualmente para cada região cerebral, e o volume de cada região também será usado na formação dos vetores de atributos de um dos grafos propostos;
- b) Apesar da busca extensiva pela melhor combinação de parâmetros, *graph kernel*, atributos de vértices, e atributo de textura, o método está restrito ao uso de somente uma forma de seleção de atributos, que é o método de limiarização de arestas

nos grafos. Esse fato se contradiz aos trabalhos apresentados na literatura, que utilizam métodos de seleção de atributos mais sofisticados a fim de obter melhores resultados de classificação.

No capítulo seguinte serão apresentados os resultados obtidos do desenvolvimento da pesquisa.

Capítulo 4

Resultados e Discussões

Neste capítulo é apresentado os resultados obtidos e as discussões do processamento das imagens, da representação das características de textura e dos experimentos de classificação.

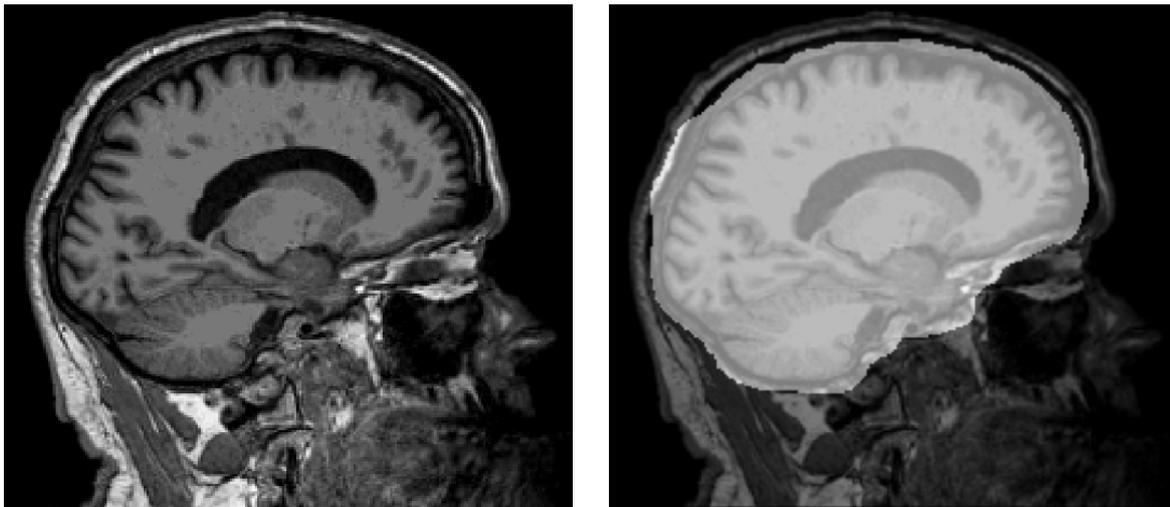
4.1 Pipeline do FreeSurfer e segmentação

As imagens de RM utilizadas neste estudo foram processadas pelo *pipeline* do FreeSurfer para que, ao fim do processamento, fosse possível extrair máscaras binárias que delimitassem as regiões segmentadas. Neste projeto extraímos 92 máscaras binárias do cérebro de cada paciente, conforme descrito nas Tabelas 3 e 4 da seção 3.3.2. Dentre as regiões segmentadas, ilustramos na Figura 8 as máscaras dos hipocampus esquerdo e direito de um dado paciente, sobrepostas na imagem registrada ao espaço do FreeSurfer, no caso, a imagem do estudo foi reamostrada de $170 \times 256 \times 256$ para $256 \times 256 \times 256$.



Figura 8 – Ilustração de duas máscaras binárias dos hipocampus esquerdo e direito extraídas e alinhadas à imagem de original do paciente, em cortes (a) Axial (b) Sagital e (c) Coronal.

Usando o FreeSurfer também extraímos para cada paciente uma máscara abrangendo todo o cérebro, a qual exclui a caixa craniana. Essa máscara tem seu uso exclusivo pelos filtros de textura GLCM, RLM e LBP, com intuito de reduzir o tempo de filtragem visto que o processamento será realizado apenas na área delimitada pela máscara. Todas as máscaras contendo o cérebro foram dilatadas em três voxels em todas as direções, a fim de corrigir possíveis erros de corte nas bordas das máscaras. O valor de três voxels foi definido após análise e teste visual usando diversas imagens.



(a) Imagem registrada para o espaço do FreeSurfer.

(b) Sobreposição da máscara binária calculada a partir da imagem (a).

Figura 9 – Cortes sagitais de uma imagem de um dado paciente após o seu registro com o espaço do FreeSurfer.

A Figura 9 ilustra um exemplo da máscara binária do cérebro para um dado paciente. Já a Figura 9(a) apresenta um corte sagital da imagem de um paciente já registrada ao espaço do FreeSurfer, e na Figura 9(b) é apresentado o mesmo corte porém com a sobreposição de sua respectiva máscara do cérebro. É possível notar que a expansão em 3 voxels surtiu efeito positivo em isolar as regiões cerebrais da borda de corte, evitando problemas no cálculo dos filtros de textura. Na Figura 10 ilustramos o mesmo corte sagital do paciente apresentado na Figura 9, porém com a sobreposição das máscaras das regiões cerebrais segmentadas pelo FreeSurfer. Nessa imagem, cada cor representa uma região cerebral a ser analisada, e todas essas regiões são extraídas na forma de máscaras binárias para delimitar a filtragem pelos algoritmos de textura. A partir da contagem de voxels das máscaras e da resolução da imagem é possível obter o volume das regiões.

4.2 Filtros de textura

Considerando que utilizamos atributos de textura que são calculados a partir dos valores de intensidade dos voxels e suas vizinhanças, isso implica que estes atributos são

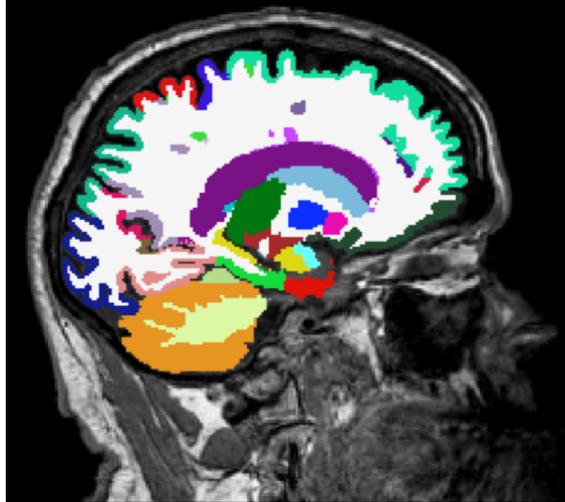
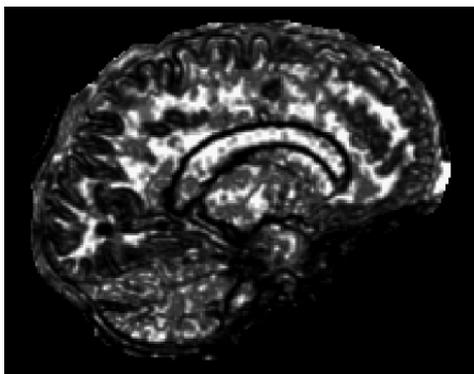
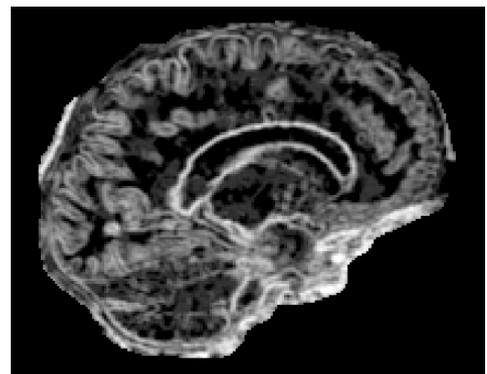


Figura 10 – Ilustração das máscaras binárias obtidas pela segmentação da imagem do paciente apresentado na Figura 9 em um corte sagital.

sensíveis à variação de intensidades dos voxels e, por esse motivo, optamos por utilizar como entrada dos filtros a imagem original extraída da base de dados da ADNI somente registrada ao espaço do FreeSurfer, isto é, sem normalização ou padronização dos histogramas de intensidade. Assim, com auxílio da máscara do cérebro, ilustrada na Figura 9, extraímos 22 atributos de textura com as técnicas GLCM, RLM e LBP. Esses atributos são inicialmente representados na forma de mapas de textura que corresponde a uma imagem filtrada.



(a) Atributo de textura F_1



(b) Atributo de textura F_2

Figura 11 – Dois atributos do GLCM em cortes sagitais de uma imagem do paciente apresentado na Figura 9.

Na Figura 11 ilustramos dois atributos de textura do GLCM, o F_1 e F_2 nas Figuras (a) e (b), respectivamente. Na Figura 12 apresentamos os atributos F_{12} e F_{16} do RLM, e na Figura 13 ilustramos os atributos F_{19} e F_{21} do LBP. Todas as figuras aqui citadas são da mesma imagem e corte do paciente apresentado na Figura 9.

A partir das Figuras 11, 12 e 13 é possível notar que as extremidades de algumas imagens possuem artefatos de filtragem pelo fato do limite da máscara possuir uma mudança

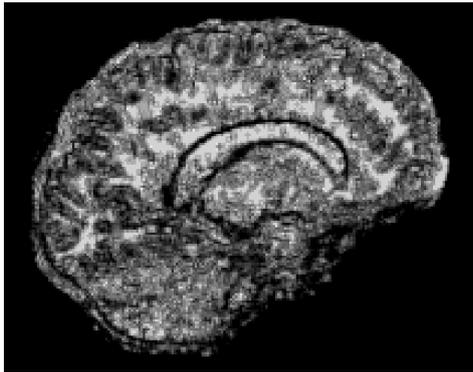
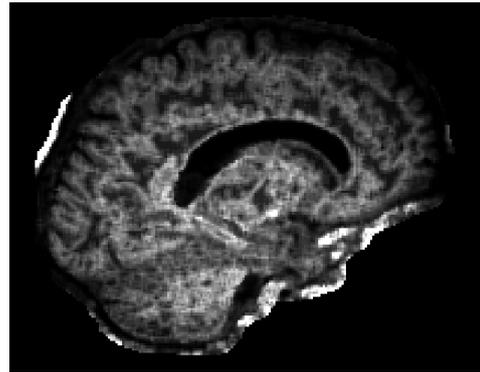
(a) Atributo de textura F_{12} (b) Atributo de textura F_{16}

Figura 12 – Dois atributos do RLM em cortes sagitais de uma imagem do paciente apresentado na Figura 9.

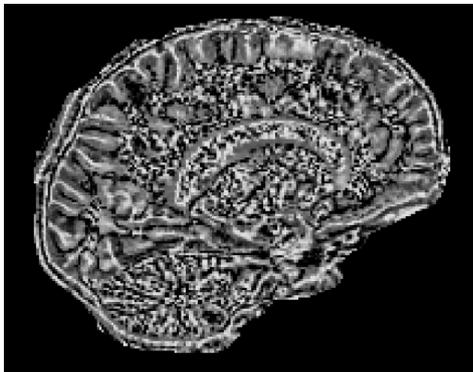
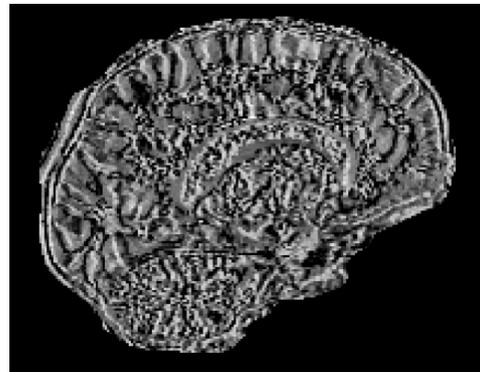
(a) Atributo de textura F_{19} (b) Atributo de textura F_{21}

Figura 13 – Dois atributos do LBP em cortes sagitais de uma imagem do paciente apresentado na Figura 9

brusca de intensidade. Esse efeito está mais visível na Figura 12(b), onde observa-se partes esbranquiçadas em algumas das extremidades da imagem. Porém, esses artefatos de filtragem não interferem nos valores de textura das regiões extraídas para a construção dos grafos, pois as regiões selecionadas estão abaixo da borda de corte da máscara cerebral.

4.3 Extração de características e normalizações

As características de textura (Seção 3.5.2) e os valores estatísticos (Seção 3.5.3) extraídos das regiões do cérebro das imagens de RM desempenham um papel importante neste estudo, pois suas distribuições compõem os atributos de vértices dos grafos. Ao aplicar as etapas de normalização, os atributos de textura, que eram separáveis entre classes, devem permanecer separáveis, e a distância entre as distribuições de textura também devem permanecer relativamente similar. Vale lembrar que as distribuições são calculadas usando um número fixo de bins, obtido pela regra de Freedman Diaconis.

A Figura 14 apresenta a distribuição de textura de três pacientes aleatórios das classes

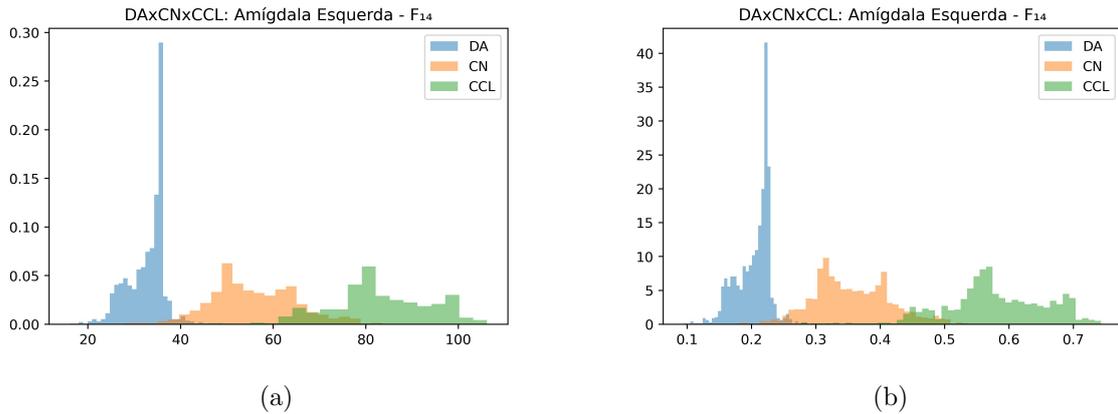


Figura 14 – Distribuições de textura (a) não normalizadas e (b) normalizadas para a região da amígdala esquerda de um paciente para cada classe diagnóstica. As distâncias de Wasserstein entre as distribuições normalizadas são $D_{ws} = 0,370$, $D_{ws} = 0,221$ e $D_{ws} = 0,148$ para os casos $CCL \times DA$, $CN \times CCL$ e $CN \times DA$, respectivamente.

CN, CCL e DA. Cada distribuição representa os valores do atributo de textura F_{14} , extraído com a técnica RLM. Essas distribuições são da amígdala esquerda de cada paciente, dado que, juntamente com os hipocampos, essa é uma das primeiras regiões a ser afetada pela progressão da DA. A hipótese nessa pesquisa é que as mudanças histopatológicas da DA, causadas pelo acúmulo de plaquetas β amiloide ($A\beta$) e entrelaçados neurofibrilares, modificam os padrões de textura das regiões relacionadas com a cognição e essas mudanças são diferentes para pacientes CN, CCL e DA.

Na Figura 14(a), cada distribuição não-normalizada representa os valores de textura do atributo extraído diretamente de seu mapa, calculado, nesse caso, pelo RLM. Nota-se também uma separação entre as três distribuições, sendo que as mesmas se encontram num intervalo aproximado entre 20 e 100. A 14(b) apresenta o resultado da aplicação das duas etapas de normalização descrito na Seção 3.5. A primeira etapa consiste na utilização da técnica min-max por atributo de textura entre todos os pacientes, a segunda aplica a técnica min-max em todas as distribuições de textura por região. Nesse caso, após a primeira etapa de normalização, o número de bins de cada região é estimado e fixado por atributo de textura.

O aumento na granularidade de bins da distribuição da Figura 14(b) comparada com a Figura 14(a) pode ser atribuído aos diferentes tamanhos das regiões e, conseqüentemente, na variabilidade do número de voxels por região. Como resultado da nossa metodologia para calcular um número fixo de bins por atributo de textura com a regra de Freedman-Diaconis, as distribuições construídas para regiões pequenas, como a amígdala, apresentaram um acréscimo no número de bins, enquanto que para regiões grandes, como o putamen, esse número se reduziu. Vale lembrar que pela nossa metodologia, o número fixo de bins é calculado a partir da média do número de bins de todas as regiões para um

dado atributo de textura.

Também é possível observar que as distâncias entre as distribuições nas Figuras 14(a) e (b) permaneceu similar, apenas com uma redução sutil na sobreposição entre as distribuições CN e CCL após as etapas de normalização. As distribuições apresentadas na Figura 14(b), também refletem sua separação pela distância de Wasserstein, onde as distribuições CCL×DA estão mais afastadas ($D_{ws} = 0.370$), seguida de CN×CCL ($D_{ws} = 0.221$) e CN×DA ($D_{ws} = 0.148$).

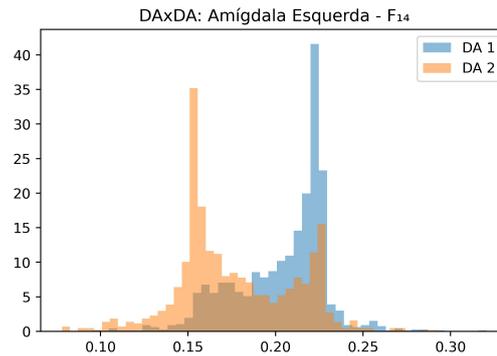


Figura 15 – Distribuições de textura para o atributo F_{14} do RLM, para dois pacientes DA (DA 1 e DA 2). A região mais escura na imagem representa a sobreposição das duas distribuições.

A Figura 15 apresenta as distribuições de textura do atributo F_{14} do RLM extraídas da amígdala esquerda de dois pacientes com DA, sendo que o paciente identificado por DA 1 refere-se ao mesmo paciente apresentado na Figura 14. Como as distribuições são do mesmo grupo de diagnóstico, podemos observar uma sobreposição significativa entre elas. A distância de Wasserstein, nesse caso, é igual a 0.027, o que é consideravelmente inferior às distâncias entre as classes apresentadas na Figura 14. Estes resultados sugerem a eficácia das etapas de normalização dos atributos de textura e da quantidade de bins, que é observada pelo fato de não alterar de forma significativa o formato das distribuições ao mesmo tempo em que mantém a separabilidade das distribuições.

4.4 Classificação e discussão

Para os experimentos de classificação, a avaliação do melhor classificador foi realizada usando a função *GridSearchCV* da biblioteca *sklearn*¹, com o parâmetro C assumindo valores dentro do conjunto $\{1e - 06, 1e - 05, 0.0001, 0.001, 0.01, 0.1, 1.0, 10.0, 100.0, 1000.0\}$ e com validação cruzada com 10 *folds*. O esquema de classificação seguiu o modelo descrito na Seção 3.8 com intuito de obter a melhor combinação de atributos de vértices, métrica de pesos, parâmetro (n) da equação de limiar e *graph kernel*, sendo que cada combinação foi executada dez vezes para então extrairmos os valores de acurácia, sensibilidade,

¹ http://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html

especificidade e AUC das dez execuções. O parâmetro do limiar (n), que multiplica o desvio padrão na equação 33, foi testado para os valores $n = 0.5, 1, 2, 3, 4, 5$. Visto que construímos grafos completos, então um valor de n maior implica em mais arestas sendo removidas. Cada linha das tabelas de resultados apresentados nesta seção, possuem a melhor combinação, em termos da maximização da acurácia, da métrica usada para o cálculo dos pesos das arestas, do atributo de textura e do valor de n .

A Tabela 5 apresenta os resultados da classificação para CN×DA, com os resultados da melhor combinação destacados em negrito. Como pode ser observado, o melhor resultado (83.8% de acurácia, 68.5% de sensibilidade, 92.7% de especificidade e 0.92 de AUC) foi alcançado usando o *graph kernel* HGK+SP para o atributo de textura F_{19} do LBP, com Vol+Stats como atributos dos vértices e o peso arestas arestas calculados a partir da distância KL. Para os outros resultados da tabela, nota-se que os valores de sensibilidade foram inferiores quando comparados à especificidade.

Graph Kernel	Tipo de vértices	Métrica das arestas	Atributo de Textura	Parâmetro do Limiar	ACC	SENS	SPEC	AUC
P2K	Distribuição	Hellinger	F_{19} (LBP)	$n = 3$	75.6 ± 2.1	46.0 ± 5.2	91.6 ± 1.6	0.72 ± 0.02
	Vol+Stats	Hellinger	F_{21} (LBP)	$n = 3$	70.2 ± 1.8	24.0 ± 4.8	95.5 ± 2.2	0.65 ± 0.02
HGK+WL	Distribuição	Hellinger	F_{21} (LBP)	$n = 2$	77.3 ± 0.6	42.0 ± 2.4	96.8 ± 0.6	0.84 ± 0.01
	Vol+Stats	Wasserstein	F_{21} (LBP)	$n = 3$	81.0 ± 1.3	52.2 ± 3.9	97.3 ± 0.6	0.87 ± 0.01
HGK+SP	Distribuição	Wasserstein	F_2 (GLCM)	$n = 2$	75.9 ± 1.7	48.3 ± 7.4	91.0 ± 2.2	0.78 ± 0.02
	Vol+Stats	KL	F_{19} (LBP)	$n = 4$	83.8 ± 1.5	68.5 ± 2.7	92.7 ± 1.7	0.92 ± 0.01
HGK+wSP	Distribuição	Wasserstein	F_{19} (LBP)	$n = 5$	73.6 ± 0.8	25.4 ± 2.8	100 ± 0.0	0.66 ± 0.01
	Vol+Stats	Hellinger	F_{19} (LBP)	$n = 2$	77.3 ± 1.6	49.5 ± 6.3	93.1 ± 1.7	0.83 ± 0.01

Tabela 5 – Resultados de classificação para CN×DA. A acurácia (ACC), sensibilidade (SENS) e especificidade (SPEC) são dadas em porcentagem.

Para o experimento CN×DA, exceto para a combinação HGK+SP+Distribuição+Wasserstein, observa-se uma prevalência dos atributos F_{19} e F_{21} do LBP entre os melhores resultados, mostrando que os atributos do LBP desempenharam melhor que os atributos do GLCM e RLM para discriminar entre CN e DA. Esses resultados também mostram que para esse experimento, os grafos com atributos de vértice do tipo Vol+Stats desempenharam melhor que os grafos utilizando a distribuição de probabilidade de textura das regiões. Isto pode ser explicado pelo fato que nos extremos da análise, ou seja, CN e DA, as diferenças de atrofia cerebral das regiões segmentadas serem mais evidentes nas imagens e que, portanto, o volume tem um peso maior.

A Tabela 6 apresenta os resultados da classificação CN×CCL usando a mesma organização dos resultados do caso CN×DA. O melhor resultado (74.1% de acurácia, 71.6% de sensibilidade, 76.2% de especificidade e 0.81 de AUC) foi obtido utilizando o *graph kernel*

Graph Kernel	Tipo de vértices	Métrica das arestas	Atributo de Textura	Parâmetro do Limiar	ACC	SENS	SPEC	AUC
P2K	Distribuição	KL	F_{13} (RLM)	$n = 0.5$	68.4 ± 2.1	59.6 ± 7.8	76.6 ± 4.2	0.75 ± 0.01
	Vol+Stats	Hellinger	F_7 (GLCM)	$n = 1$	64.9 ± 1.8	50.2 ± 7.3	77.8 ± 3.7	0.70 ± 0.01
HGK+WL	Distribuição	Hellinger	F_{17} (RLM)	$n = 3$	68.8 ± 1.4	57.3 ± 6.2	79.0 ± 3.5	0.78 ± 0.01
	Vol+Stats	Hellinger	F_{17} (RLM)	$n = 3$	74.1 ± 1.3	71.6 ± 5.0	76.2 ± 2.4	0.81 ± 0.01
HGK+SP	Distribuição	Hellinger	F_{17} (RLM)	$n = 3$	68.5 ± 2.9	63.0 ± 7.1	74.2 ± 2.3	0.77 ± 0.01
	Vol+Stats	KL	F_{18} (RLM)	$n = 1$	74.0 ± 1.0	79.8 ± 2.3	68.9 ± 2.6	0.81 ± 0.01
HGK+wSP	Distribuição	KL	F_{17} (RLM)	$n = 0.5$	64.8 ± 1.8	51.0 ± 6.7	78.1 ± 3.0	0.69 ± 0.01
	Vol+Stats	Wasserstein	F_5 (GLCM)	$n = 5$	70.4 ± 2.2	66.1 ± 6.9	75.7 ± 2.6	0.79 ± 0.01

Tabela 6 – Resultados de classificação para $CN \times CCL$. A acurácia (ACC), sensibilidade (SENS) e especificidade (SPEC) são dadas em porcentagem.

HGK+WL para o atributo de textura F_{17} do RLM, com Vol+Stats como atributos dos vértices e os pesos das arestas calculadas pela distância de Hellinger. Similarmente aos resultados de $CN \times DA$, os resultados de acurácia para os atributos de vértices Vol+Stats foram superiores aos resultados com a distribuição dos atributos de textura das regiões. Entretanto, em contraste aos resultados de $CN \times DA$, as diferenças entre sensibilidade e especificidade foram menores. Ao analisar a coluna de atributos de textura, nota-se que seis dos oito resultados foram obtidos para os atributos de textura do RLM. Além disso, também é válido salientar que a combinação HGK+SP com Vol+Stats resultou no mesmo valor de AUC (0.81) e com uma acurácia levemente inferior (74.0%).

Graph Kernel	Tipo de vértices	Métrica das arestas	Atributo de Textura	Parâmetro do Limiar	ACC	SENS	SPEC	AUC
P2K	Distribuição	Hellinger	F_{14} (RLM)	$n = 2$	70.7 ± 1.7	33.0 ± 3.9	94.1 ± 1.9	0.67 ± 0.01
	Vol+Stats	Wasserstein	F_{22} (LBP)	$n = 3$	68.0 ± 1.1	19.5 ± 3.9	98.8 ± 0.9	0.63 ± 0.01
HGK+WL	Distribuição	Wasserstein	F_{12} (RLM)	$n = 2$	74.1 ± 1.3	38.7 ± 3.2	96.4 ± 0.5	0.68 ± 0.01
	Vol+Stats	KL	F_{15} (RLM)	$n = 3$	74.0 ± 2.5	41.1 ± 3.2	94.3 ± 1.3	0.72 ± 0.02
HGK+SP	Distribuição	Wasserstein	F_{15} (RLM)	$n = 5$	74.8 ± 1.3	41.3 ± 2.0	95.4 ± 2.4	0.72 ± 0.02
	Vol+Stats	KL	F_{13} (RLM)	$n = 3$	75.4 ± 1.3	50.0 ± 3.2	91.0 ± 2.4	0.74 ± 0.01
HGK+wSP	Distribuição	Hellinger	F_{22} (LBP)	$n = 5$	71.9 ± 1.9	34.5 ± 5.3	96.3 ± 0.8	0.73 ± 0.01
	Vol+Stats	Hellinger	F_{20} (LBP)	$n = 3$	70.9 ± 0.6	24.9 ± 2.3	99.9 ± 0.2	0.71 ± 0.01

Tabela 7 – Resultados de classificação para $CCL \times DA$. A acurácia (ACC), sensibilidade (SENS) e especificidade (SPEC) são dadas em porcentagem.

Para os resultados de $CCL \times DA$ apresentados na Tabela 7, o melhor resultado (75.4% de acurácia, 50% de sensibilidade, 91% de especificidade e 0.74 de AUC) foi alcançado pela combinação do *graph kernel* HGK+SP com o atributo de textura F_{13} do RLM, utilizando o volume e os momentos estatísticos como atributos dos vértices, e o pesos das arestas

calculados pela distância KL. Essa combinação foi a mesma usada para obter o melhor resultado em CN×DA. Os melhores resultados de classificação no experimento CCL×DA foram obtidos para altos valores do parâmetro do limiar (n) quando comparado com os casos CN×DA e CN×CCL. Isso é um indicativo que para o caso CCL×DA foi necessário remover mais arestas para obter grafos discriminantes entre as classes de diagnóstico. Similarmente ao caso CN×DA, os valores de sensibilidade foram baixos.

Uma observação que vale ressaltar é que nenhum dos oito atributos do GLCM apareceu entre os melhores resultados no experimento de CCL×DA. Além disso, comparando os três casos de classificação (CN×DA, CN×CCL e CCL×DA), observa-se que os atributos de textura do RLM e do LBP foram mais discriminantes para os casos envolvendo o grupo AD, isto é, para os casos CN×DA e CCL×DA.

4.4.1 Comparação com métodos da literatura

Métodos	Pacientes			CN×DA	CN×CCL	CCL×DA
	CN	CCL	DA	ACCAUC	ACCAUC	ACCAUC
(LIU et al., 2017)	230	280	200	95.3 0.96	86.6 0.91	86.5 0.85
(JIE et al., 2018)	50	99	34	– –	82.6 0.78	– –
(CUI et al., 2018)	21	25	22	91.3 0.99	98.3 0.99	77.3 0.97
(WEE et al., 2019)	242	415	355	81.0 –	67.7 –	65.4 –
(HETT et al., 2021)	213	–	130	91.6 –	– –	– –
Método proposto	200	153	121	83.8 0.92	74.1 0.81	75.4 0.74

Tabela 8 – Comparação de resultados com outros métodos. A acurácia (ACC) é dada em porcentagem.

Para fins de comparação, a Tabela 8 apresenta os valores de acurácia e AUC do método apresentado nesta pesquisa (última linha) e de outros cinco métodos publicados na literatura, sendo que os estudos comparados utilizaram abordagens baseadas em grafos e usando imagens da base ADNI. Uma breve descrição desses métodos é apresentada a seguir:

- Liu et al. (2017) apresentou um método em grafos construídos a partir de múltiplas regiões com atributos do GLCM e realizando duas etapas de seleção de atributos nos grafos. Embora em geral seus resultados sejam superiores aos obtidos pelo nosso método, para o caso CN×AD o valor da AUC do nosso método é apenas ligeiramente inferior. No entanto, é importante notar que em seu trabalho, Liu et al. não forneceu os valores de variância das métricas utilizadas para avaliar os resultados.

- Jie et al. (2018) propuseram uma abordagem baseada em *graph kernel* a partir de múltiplas regiões cerebrais. Os autores utilizaram imagens de fMRI da ADNI para a extração de atributos. Em comparação com este método, nossa abordagem alcançou um valor de AUC superior para o caso CN×CCL, ou seja, 0,81 de nosso método contra 0,78 da abordagem dos autores. O trabalho deles também usou menos imagens quando comparado ao nosso estudo.
- Cui et al. (2018) propuseram um método baseado em *graph kernel* que utilizou como atributos as leituras extraídas de imagens de fMRI de múltiplas regiões cerebrais. Apesar dos autores terem relatado resultados superiores, eles usaram menos de 30 imagens para cada classe diagnóstica.
- Wee et al. (2019) propuseram um método baseado em um *spectral graph convolutional neural network* que incorpora a espessura cortical das regiões e suas informações de geometria subjacente para identificar CCL e DA usando dados de RM T1-w da ADNI. Embora o foco dos autores tenha sido na predição da conversão de CCL em DA, eles também usaram o método para a classificação das imagens em pares de classes diagnósticas. Em comparação com o trabalho desses autores, o nosso método resultou em uma acurácia superior em todos os casos de classificação. No entanto, vale ressaltar que o número de imagens utilizadas em nosso estudo foi inferior para as classes CCL e AD.
- Hett et al. (2021) usaram múltiplas regiões cerebrais e extraíram seus atributos com o método PBG (COUPÉ et al., 2012). O foco do estudo foi principalmente distinguir CCLe de CCLpr, mas eles também usaram o método para o caso CN×AD, onde obtiveram resultados um pouco acima dos nossos com número semelhante de pacientes.

Em contraste com o nosso método, que usa apenas um método baseado em limiares para remover arestas dos grafos, os métodos comparados acima usam estratégias sofisticadas de seleção de atributos de arestas e vértices para construir grafos discriminantes entre as classes de diagnóstico, o que pode ter contribuído positivamente para os resultados da classificação. Também é válido salientar que os trabalhos nem todos os trabalhos apresentados são diretamente comparáveis ao nosso, pois não utilizaram imagens de RM estrutural e atributos de textura nos grafos. Dado que o uso de RM estrutural é positivo do ponto de vista clínico, por não necessitar do uso do contraste endovenoso no paciente, o que representa um custo a menos para a extração da imagem de RM, assim como, positivo para a saúde do paciente em tratamento.

4.5 Considerações finais

Este capítulo apresentou os resultados obtidos do processamento inicial das imagens de RM, extração de máscaras de regiões cerebrais, e da aplicação dos filtros de textura. Também foi discutido os efeitos das etapas de normalização e sobre o tratamento dos dados de textura e suas distribuições de probabilidade, assim como os resultados de classificação dos grafos e as comparações com o estado da arte envolvendo classificação de imagens da ADNI com métodos baseados em grafos.

Avaliamos os resultados da classificação de grafos a partir dos *graph kernels* P2K e HGK, sendo que o HGK foi utilizado com três *kernels* base para uma melhor investigação de qual deles resultaria em melhores resultados por classe de diagnóstico. Também foi mostrado que uma abordagem por atributos de textura extraídos de múltiplas regiões cerebrais é viável, dado que este foi comparado a outros métodos da literatura. Nesta pesquisa, avaliamos 22 atributos de textura extraídos com três das principais técnicas de extração de atributos de textura, o GLCM, RLM e LBP. Pelos resultados de classificação do caso CN×DA, foi possível verificar que independente do *graph kernel* utilizado e do tipo de atributos de vértice, os atributos de textura do LBP foram ranqueados de forma superior aos atributos das outras duas técnicas.

Como uma das considerações sobre os resultados, vale ressaltar que para as classificações CN×DA e CCL×DA, isto é, os casos envolvendo o grupo de diagnóstico com DA, os valores de sensibilidade obtidos foram menores em relação aos da especificidade. Esse problema talvez possa ser explicado pelo desbalanço do número de imagens utilizadas pelo grupo DA quando comparado ao grupo CN e CCL, sendo este um ponto a ser melhorado em futuros trabalhos.

Avaliando os resultados de classificação entre os três casos apresentados, é possível notar que uma técnica de extração de textura prevalece nos melhores resultados das combinações. Isso pode ser um ponto de partida para sanar uma necessidade deste projeto, que é falta de um método para seleção de atributos. Sejam estes atributos como os atributos dos vértices, bem como, a possibilidade de mesclar e selecionar atributos de uma mesma técnica para compor o vetor de atributos das regiões em seus grafos. Isto é, focando nos atributos da técnica que prevaleceu como melhor discriminante entre as classes classificadas.

Capítulo 5

Conclusões

Este capítulo apresenta as conclusões finais do trabalho, enumerando suas contribuições, ideais para trabalhos futuros e a produção científica desenvolvida no decorrer da pesquisa.

5.1 Conclusões

Nesta pesquisa propusemos o desenvolvimento de uma técnica computacional baseada em *graph kernel* para a classificação de grafos cerebrais construídos a partir de imagens de RM estruturais, com a extração de atributos de textura de múltiplas regiões cerebrais. A motivação por trás dessa pesquisa está no crescente número de casos da doença de Alzheimer e, como consequência, na demanda por técnicas computacionais automáticas que consigam auxiliar no diagnóstico da doença por imagens de ressonância magnética, que já são rotineiras no tratamento e diagnóstico clínico do Alzheimer. Também pelo recente movimento na literatura para o desenvolvimento de novos biomarcadores para a DA, como a investigação de danos causados aos tecidos que possam ser identificados a partir de algoritmos de textura.

A metodologia foi baseada na segmentação das imagens de RM em múltiplas regiões cerebrais, das quais foram extraídos 22 atributos de textura com os filtros GLCM, RLM e LBP. Nesse sentido, desenvolvemos duas abordagens para trabalhar com esses atributos, a primeira usando a distribuição de probabilidade dos atributos de textura normalizados e segundo, por um vetor contendo o volume e os quatro primeiros momentos estatísticos de cada região segmentada. Avaliamos, por meio de um esquema de classificação, todas as combinações de quatro *graph kernels*, 22 atributos de textura, duas abordagens para o vetor de atributos dos vértices dos grafos. Além disso, ainda avaliamos as distâncias de Wasserstein, Hellinger e KL para o peso das arestas, que foi calculado entre as distribuições de probabilidade de textura das regiões adjacentes. Como uma forma de seleção de arestas

deses grafos, desenvolvemos um método baseado em limiar, capaz de trabalhar com as três distâncias para cada atributo de textura.

Os resultados de classificação mostraram a eficácia da nossa metodologia, ao se mostrarem comparáveis aos resultados de métodos recentes da literatura. Os melhores de resultados de classificação apresentaram os seguintes valores de AUCs: $AUC = 0.92$ para o caso $CN \times DA$, $AUC = 0.81$ para $CN \times CCL$ e $AUC = 0.78$ para $CCL \times DA$. Também foi mostrado que para o caso $CN \times DA$ a abordagem de atributos dos vértices que relacionavam o volume da região com os quatro primeiros momentos estatísticos foram superiores a outra abordagem que utilizava a distribuição de probabilidade de atributos de textura da região. Esse dado pode ser explicado pela maior diferença de volume entre as classes CN e DA causada pela atrofia progressiva. Dado a variedade de atributos de textura avaliados nesta pesquisa, pudemos constatar que os atributos do LBP foram mais discriminantes para $CN \times DA$, os do RLM foram mais discriminantes para o caso $CN \times CCL$ e um balanço de atributos RLM e LBP para o caso $CCL \times DA$. Estes dados podem sugerir que os atributos de textura do LBP atuam melhor com imagens de pacientes com DA .

Em comparação aos métodos do estado da arte, o nosso trabalho não incluiu métodos para seleção de atributos de textura e métodos para seleção de vértices nos grafos. Nesta pesquisa, somente foi utilizado um método para seleção das arestas mais discriminantes entre os grafos. Sendo este, um ponto de futura investigação para trabalhos futuros, como o impacto nos resultados ao aplicar métodos de seleção de atributos.

5.2 Trabalhos futuros

Entre as ideias para trabalhos futuros, podemos citar a inclusão de outros atributos à metodologia, tal como no trabalho de Hett et al. (2021) que também utilizou as pontuações de testes cognitivos como o MMSE, e também incluiu em *ensemble* um grafo das sub-regiões do hipocampo, conforme mencionamos na Seção 3.3.1. Outra sugestão de melhoria seria a inclusão de métodos para seleção de atributos de textura, e também da simplificação dos 22 atributos de textura como um vetor contendo a média do atributo para a região. Também vale ressaltar o aumento da base de dados utilizada, com intuito de diminuir ou sanar o desbalanceio entre classes.

5.3 Contribuições

As principais contribuições desta pesquisa, foram:

- O desenvolvimento de um novo método baseado em *graph kernel* que permite o uso de vários atributos de textura de imagens 3D;
- A avaliação de três filtros de textura e dois métodos de extração de características aplicados a regiões cerebrais segmentadas e a demonstração de suas capacidades

de classificação;

- O desenvolvimento de um método baseado em limiar para remover arestas de grafos completos com o intuito de selecionar o modelo mais discriminativo;
- A avaliação de três distâncias diferentes nos pesos das arestas para comparar distribuições de probabilidade;
- A análise exaustiva do método proposto usando um número significativo de imagens de RM.

5.4 Trabalho submetido

Como resultado desta pesquisa, um artigo foi submetido em Dezembro de 2021 para a revista *Expert Systems with Applications*¹ com o título “*Alzheimer’s disease classification based on graph kernel SVMs constructed with 3D texture features extracted from MR images*”. No momento da escrita deste texto, o artigo encontra-se em revisão.

¹ <https://www.journals.elsevier.com/expert-systems-with-applications>

Referências

- ALZHEIMER'S ASSOCIATION. Advancing alzheimer's research worldwide. Chigaco, EUA, 2020. Disponível em: <https://www.alz.org/brain_portuguese/09.asp>.
- ALZHEIMER'S DISEASE INTERNATIONAL. The global voice on dementia. Reino Unido, 2020. Disponível em: <<https://www.alz.co.uk/>>.
- BANERJEE, J. et al. 3d lbp-based rotationally invariant region description. In: SPRINGER. **Asian Conference on Computer Vision**. [S.l.], 2012. p. 26–37.
- BEHESHTI, I. et al. Histogram-Based Feature Extraction from Individual Gray Matter Similarity-Matrix for Alzheimer's Disease Classification. **Journal of Alzheimer's Disease**, v. 55, n. 4, p. 1571–1582, 2017. ISSN 18758908.
- BISHOP, C. M. **Pattern recognition and machine learning**. [S.l.]: springer, 2006.
- BORGWARDT, K. M.; KRIEGEL, H.-P. Shortest-path kernels on graphs. In: IEEE. **Fifth IEEE International Conference on Data Mining (ICDM'05)**. [S.l.], 2005. p. 1–8.
- BUSHBERG, J. T.; BOONE, J. M. **The essential physics of medical imaging**. [S.l.]: Lippincott Williams & Wilkins, 2011.
- BUSTAMAM, A.; SARWINDA, D.; ARDENASWARI, G. Texture and gene expression analysis of the MRI brain in detection of Alzheimer's disease. **Journal of Artificial Intelligence and Soft Computing Research**, v. 8, n. 2, p. 111–120, 2018.
- BUTTERFIELD, D. A.; HALLIWELL, B. Oxidative stress, dysfunctional glucose metabolism and Alzheimer disease. **Nature Reviews Neuroscience**, v. 20, n. 3, p. 148–160, 2019.
- CAI, J. H. et al. Magnetic Resonance Texture Analysis in Alzheimer's disease. **Academic Radiology**, v. 27, n. 12, p. 1774–1783, 2020.
- CASTELLANO, G. et al. Texture analysis of medical images. **Clinical radiology**, Elsevier, v. 59, n. 12, p. 1061–1069, 2004.
- CAUCHY, A. L. Sur les résultats moyens d'observations de même nature, et sur les résultats les plus probables. **CR Acad. Sci. Paris**, v. 37, p. 198–206, 1853.

COLLINS, D. L. et al. Automatic 3d intersubject registration of mr volumetric data in standardized talairach space. **Journal of computer assisted tomography**, v. 18, n. 2, p. 192–205, 1994.

CORTES, C.; VAPNIK, V. Support-vector networks. **Machine learning**, Springer, v. 20, n. 3, p. 273–297, 1995.

COUPÉ, P. et al. NeuroImage : Clinical Scoring by nonlocal image patch estimator for early detection of Alzheimer ' s disease. **YNICL**, The Authors, v. 1, n. 1, p. 141–152, 2012. ISSN 2213-1582. Disponível em: <<http://dx.doi.org/10.1016/j.nicl.2012.10.002>>.

COUPÉ, P. et al. Scoring by nonlocal image patch estimator for early detection of alzheimer's disease. **NeuroImage: clinical**, Elsevier, v. 1, n. 1, p. 141–152, 2012.

CUI, X. et al. Classification of alzheimer's disease, mild cognitive impairment, and normal controls with subnetwork selection and graph kernel principal component analysis based on minimum spanning tree brain functional network. **Frontiers in Computational Neuroscience**, v. 12, n. May, p. 1–12, 2018. ISSN 16625188.

DATAR, M. et al. Locality-sensitive hashing scheme based on p-stable distributions. In: **Proceedings of the twentieth annual symposium on Computational geometry**. [S.l.: s.n.], 2004. p. 253–262.

DESIKAN, R. S. et al. An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest. **NeuroImage**, v. 31, n. 3, p. 968 – 980, 2006. ISSN 1053-8119. Disponível em: <<http://www.sciencedirect.com/science/article/B6WNP-4JFHF4P-1/2/0ec667d4c17eafb0a7c52fa3fd5aef1c>>.

DESTRIEUX, C. et al. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. **NeuroImage**, v. 53, n. 1, p. 1–15, 2010. ISSN 1053-8119. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1053811910008542>>.

DIECKMEYER, M. et al. Texture features of proton density fat fraction maps from chemical shift encoding-based mri predict paraspinal muscle strength. **Diagnostics**, Multidisciplinary Digital Publishing Institute, v. 11, n. 2, p. 239, 2021.

DUBOIS, B. et al. Advancing research diagnostic criteria for alzheimer's disease: the iw-g-2 criteria. **The Lancet Neurology**, Elsevier, v. 13, n. 6, p. 614–629, 2014.

ENKIRCH, S. J. et al. The erica score: An mr imaging-based visual scoring system for the assessment of entorhinal cortex atrophy in alzheimer disease. **Radiology**, Radiological Society of North America, v. 288, n. 1, p. 226–333, 2018.

FENG, J.; ZHANG, S. W.; CHEN, L. Identification of Alzheimer's disease based on wavelet transformation energy feature of the structural MRI image and NN classifier. **Artificial Intelligence in Medicine**, v. 108, 2020.

FISCHL, B. et al. Automatically Parcellating the Human Cerebral Cortex. **Cerebral Cortex**, v. 14, n. 1, p. 11–22, 2004. Disponível em: <<http://cercor.oxfordjournals.org/content/14/1/11.abstract>>.

_____. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. **Neuron**, v. 33, p. 341–355, 2002.

- FISHER, R. A. Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. **Biometrika**, JSTOR, v. 10, n. 4, p. 507–521, 1915.
- FREEDMAN, D.; DIACONIS, P. On the histogram as a density estimator: L² theory. **Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete**, Springer, v. 57, n. 4, p. 453–476, 1981.
- FRENZEL, S. et al. A biomarker for alzheimer’s disease based on patterns of regional brain atrophy. **Frontiers in psychiatry**, Frontiers, v. 10, p. 953, 2020.
- FRISONI, G. B. et al. The clinical use of structural mri in alzheimer disease. **Nature Reviews Neurology**, Nature Publishing Group, v. 6, n. 2, p. 67–77, 2010.
- GALLOWAY, M. M. Texture analysis using grey level run lengths. **NASA STI/Recon Technical Report N**, v. 75, p. 18555, 1974.
- GIRAUD, R. et al. An optimized patchmatch for multi-scale and multi-feature label fusion. **NeuroImage**, Elsevier, v. 124, p. 770–782, 2016.
- GUADALUPE, T. et al. Measurement and genetics of human subcortical and hippocampal asymmetries in large datasets. **Human brain mapping**, Wiley Online Library, v. 35, n. 7, p. 3277–3289, 2014.
- GUO, H. et al. Alzheimer classification using a minimum spanning tree of high-order functional network on fmri dataset. **Frontiers in neuroscience**, Frontiers, v. 11, p. 639, 2017.
- HAN, J.; KAMBER, M.; PEI, J. **Data mining concepts and techniques, third edition**. Waltham, Mass.: Morgan Kaufmann Publishers, 2012. Disponível em: <http://www.amazon.de/Data-Mining-Concepts-Techniques-Management/dp/0123814790/ref=tmm_hrd_title_0?ie=UTF8&qid=1366039033&sr=1-1>.
- HANSEN, N. et al. Hippocampal and hippocampal-subfield volumes from early-onset major depression and bipolar disorder to cognitive decline. **Frontiers in aging neuroscience**, Frontiers, v. 13, p. 153, 2021.
- HANSON, L. G. Introduction to magnetic resonance imaging techniques. 2009.
- HARALICK, R. M. Textural Features for Image Classification. n. January 1973, 1973.
- HECKERS, S.; KONRADI, C. Hippocampal pathology in schizophrenia. In: **Behavioral neurobiology of schizophrenia and its treatment**. [S.l.]: Springer, 2010. p. 529–553.
- HELLINGER, E. Neue begründung der theorie quadratischer formen von unendlichvielen veränderlichen. **Journal für die reine und angewandte Mathematik**, De Gruyter Berlin, New York, v. 1909, n. 136, p. 210–271, 1909.
- HETT, K. et al. Graph of brain structures grading for early detection of Alzheimer ’ s disease To cite this version : HAL Id : hal-01830896 Graph of brain structures grading for early detection of Alzheimer ’ s disease. 2018.
- _____. Multi-scale graph-based grading for alzheimer’s disease prediction. **Medical Image Analysis**, Elsevier, v. 67, p. 101850, 2021.

- IGLESIAS, J. E. et al. A computational atlas of the hippocampal formation using ex vivo, ultra-high resolution mri: application to adaptive segmentation of in vivo mri. **Neuroimage**, Elsevier, v. 115, p. 117–137, 2015.
- JENKINSON, M. et al. Improved optimization for the robust and accurate linear registration and motion correction of brain images. **Neuroimage**, Elsevier, v. 17, n. 2, p. 825–841, 2002.
- JENKINSON, M.; SMITH, S. A global optimisation method for robust affine registration of brain images. **Medical image analysis**, Elsevier, v. 5, n. 2, p. 143–156, 2001.
- JIE, B. et al. Sub-network based kernels for brain network classification. **ACM-BCB 2016 - 7th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics**, p. 622–629, 2016.
- _____. Sub-Network Kernels for Measuring Similarity of Brain Connectivity Networks in Disease Diagnosis. **IEEE Transactions on Image Processing**, v. 27, n. 5, p. 2340–2353, 2018.
- JR, C. R. J. et al. The alzheimer’s disease neuroimaging initiative (adni): Mri methods. **Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine**, Wiley Online Library, v. 27, n. 4, p. 685–691, 2008.
- KLEIN, A.; TOURVILLE, J. 101 labeled brain images and a consistent human cortical labeling protocol. **Frontiers in neuroscience**, Frontiers, v. 6, p. 171, 2012.
- KRIEGE, N. M.; JOHANSSON, F. D.; MORRIS, C. A survey on graph kernels. **Applied Network Science**, v. 5, n. 1, 2020. ISSN 23648228.
- _____. **Open Access A survey on graph kernels**. [S.l.]: Applied Network Science, 2020. ISBN 4110901901.
- KRUSKAL, J. B. J., 1956. on the shortest spanning subtree of a graph and the traveling salesman problem. **Proceedings of the American Mathematical Society**, p. 48–50, 1956.
- KULLBACK, S.; LEIBLER, R. On information and sufficiencyannals of mathematical statistics, 22, 79–86. **MathSciNet MATH**, 1951.
- LEE, U.; KIM, S.; JUNG, K.-Y. Classification of epilepsy types through global network analysis of scalp electroencephalograms. **Physical Review E**, APS, v. 73, n. 4, p. 041920, 2006.
- LIU, J. et al. Classification of alzheimer’s disease using whole brain hierarchical network. **IEEE/ACM transactions on computational biology and bioinformatics**, IEEE, v. 15, n. 2, p. 624–632, 2016.
- _____. Alzheimer’s disease classification based on individual hierarchical networks constructed with 3-d texture features. **IEEE transactions on nanobioscience**, IEEE, v. 16, n. 6, p. 428–437, 2017.
- LORENA, A. C.; CARVALHO, A. C. de. Uma introdução às support vector machines. **Revista de Informática Teórica e Aplicada**, v. 14, n. 2, p. 43–67, 2007.

MA, X.; WU, G.; KIM, W. H. ENRICHING STATISTICAL INFERENCES ON BRAIN CONNECTIVITY FOR ALZHEIMER ' S DISEASE ANALYSIS VIA LATENT SPACE GRAPH EMBEDDING Department of Computer Science and Engineering , University of Texas at Arlington Department of Psychiatry , University of North Car. p. 1685–1689, 2020.

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.

MANUSCRIPT, A. ce pte d M us A novel approach to brain connectivity reveals early. 2018.

MARTINS, A. F. et al. Nonextensive information theoretic kernels on measures. **Journal of Machine Learning Research**, v. 10, n. 4, 2009.

MARTINS, R. N. et al. Alzheimers disease: a journey from amyloid peptides and oxidative stress, to biomarker technologies and disease prevention strategies—gains from AIBL and DIAN cohort studies. **Journal of Alzheimer's disease**, v. 62, n. 3, p. 965–992, 2018.

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.

MCKHANN, G. M. et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. **Alzheimer's & dementia**, v. 7, n. 3, p. 263–269, 2011.

MORRIS, C. et al. Faster kernels for graphs with continuous attributes via hashing. In: IEEE. **IEEE 16th International Conference on Data Mining (ICDM)**. [S.l.], 2016. p. 1095–1100.

MORRIS, J. C. et al. Mild cognitive impairment represents early-stage Alzheimer disease. **Archives of neurology**, v. 58, n. 3, p. 397–405, 2001.

NANNI, L. et al. Texture descriptors and voxels for the early diagnosis of Alzheimer's disease. **Artificial Intelligence in Medicine**, v. 97, p. 19–26, 2019.

NEUMANN, M. et al. Propagation kernels: efficient graph kernels from propagated information. **Machine Learning**, Springer, v. 102, n. 2, p. 209–245, 2016.

NOORTMAN, W. A. et al. Adding the temporal domain to pet radiomic features. **PLoS one**, Public Library of Science San Francisco, CA USA, v. 15, n. 9, p. e0239438, 2020.

OJALA, T.; PIETIKÄINEN, M.; HARWOOD, D. A comparative study of texture measures with classification based on featured distributions. **Pattern recognition**, Elsevier, v. 29, n. 1, p. 51–59, 1996.

PRINCE, M. et al. The global prevalence of dementia: a systematic review and metaanalysis. **Alzheimer's & dementia**, Elsevier, v. 9, n. 1, p. 63–75, 2013.

- RAMDAS, A.; TRILLOS, N. G.; CUTURI, M. On wasserstein two-sample testing and related families of nonparametric tests. **Entropy**, v. 19, n. 2, p. 1–18, 2017. ISSN 10994300.
- RICHIARDI, J. et al. Machine learning with brain graphs: Predictive modeling approaches for functional imaging in systems neuroscience. **IEEE Signal Processing Magazine**, v. 30, n. 3, p. 58–70, 2013. ISSN 10535888.
- _____. Machine learning with brain graphs: Predictive modeling approaches for functional imaging in systems neuroscience. **IEEE Signal Processing Magazine**, v. 30, n. 3, p. 58–70, 2013. ISSN 10535888.
- RUBINOV, M.; SPORNS, O. Complex network measures of brain connectivity: uses and interpretations. **Neuroimage**, Elsevier, v. 52, n. 3, p. 1059–1069, 2010.
- RUBNER, Y.; TOMASI, C.; GUIBAS, L. J. The earth mover’s distance as a metric for image retrieval. **International journal of computer vision**, Springer, v. 40, n. 2, p. 99–121, 2000.
- SCHÖLKOPF, B.; SMOLA, A.; MÜLLER, K.-R. Kernel principal component analysis. In: SPRINGER. **International conference on artificial neural networks**. [S.l.], 1997. p. 583–588.
- SCHÖLKOPF, B. et al. **Learning with kernels: support vector machines, regularization, optimization, and beyond**. [S.l.]: MIT press, 2002.
- SEGONNE, F. et al. A hybrid approach to the skull stripping problem in mri. **NeuroImage**, v. 22, n. 3, p. 1060 – 1075, 2004. ISSN 1053-8119. Disponível em: <<http://www.sciencedirect.com/science/article/B6WNP-4CF5CNY-1/2/33cc73136f06f019b2c11023e7a95341>>.
- SHARAEV, M. et al. Learning connectivity patterns via graph Kernels for fMRI-based depression diagnostics. **IEEE International Conference on Data Mining Workshops, ICDMW**, v. 2018-Novem, n. September, p. 308–314, 2019. ISSN 23759259.
- SHERVASHIDZE, N. et al. Weisfeiler-lehman graph kernels. **Journal of Machine Learning Research**, v. 12, n. 9, 2011.
- 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.
- SMITH, S. M. Fast robust automated brain extraction. **Human brain mapping**, Wiley Online Library, v. 17, n. 3, p. 143–155, 2002.
- SØRENSEN, L. et al. Early detection of Alzheimer’s disease using MRI hippocampal texture. **Human brain mapping**, v. 37, n. 3, p. 1148–1161, 2016.
- SPORNS, O. Contributions and challenges for network models in cognitive neuroscience. **Nature neuroscience**, Nature Publishing Group, v. 17, n. 5, p. 652, 2014.
- STURGES, H. A. The choice of a class interval. **Journal of the american statistical association**, New York, v. 21, n. 153, p. 65–66, 1926.

- TUSTISON, N. J. et al. N4itk: improved n3 bias correction. **IEEE transactions on medical imaging**, IEEE, v. 29, n. 6, p. 1310–1320, 2010.
- TZOURIO-MAZOYER, N. et al. Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain. **Neuroimage**, Elsevier, v. 15, n. 1, p. 273–289, 2002.
- WANG, H. et al. A learning-based wrapper method to correct systematic errors in automatic image segmentation: consistently improved performance in hippocampus, cortex and brain segmentation. **NeuroImage**, Elsevier, v. 55, n. 3, p. 968–985, 2011.
- WANG, J.; WILSON, R. C.; HANCOCK, E. R. fmri activation network analysis using bose-einstein entropy. In: SPRINGER. **Joint IAPR International Workshops on Statistical Techniques in Pattern Recognition (SPR) and Structural and Syntactic Pattern Recognition (SSPR)**. [S.l.], 2016. p. 218–228.
- WEE, C.-Y. et al. Cortical graph neural network for AD and MCI diagnosis and transfer learning across populations. **NeuroImage: Clinical**, v. 23, n. 101929, p. 1–15, 2019.
- _____. Disrupted brain functional network in internet addiction disorder: a resting-state functional magnetic resonance imaging study. **PLoS one**, Public Library of Science, v. 9, n. 9, p. e107306, 2014.
- WOOLARD, A. A.; HECKERS, S. Anatomical and functional correlates of human hippocampal volume asymmetry. **Psychiatry Research: Neuroimaging**, Elsevier, v. 201, n. 1, p. 48–53, 2012.
- XIA, H.; HOI, S. C. Mkboost: A framework of multiple kernel boosting. **IEEE Transactions on knowledge and data engineering**, IEEE, v. 25, n. 7, p. 1574–1586, 2012.
- YAAKUB, S. N. et al. On brain atlas choice and automatic segmentation methods : a comparison of MAPER & FreeSurfer using three atlas databases. **Scientific Reports**, Springer US, p. 1–15, 2020. ISSN 2045-2322. Disponível em: <<http://dx.doi.org/10.1038/s41598-020-57951-6>>.
- YAN, X.; HAN, J. gspan: Graph-based substructure pattern mining. In: IEEE. **2002 IEEE International Conference on Data Mining, 2002. Proceedings**. [S.l.], 2002. p. 721–724.
- ZHANG, J. et al. Alzheimer’s disease diagnosis using landmark-based features from longitudinal structural mr images. **IEEE journal of biomedical and health informatics**, IEEE, v. 21, n. 6, p. 1607–1616, 2017.
- _____. 3d texture analysis on mri images of alzheimer’s disease. **Brain imaging and behavior**, Springer, v. 6, n. 1, p. 61–69, 2012.
- ZHANG, Y.; BRADY, M.; SMITH, S. Segmentation of brain mr images through a hidden markov random field model and the expectation-maximization algorithm. **IEEE transactions on medical imaging**, Ieee, v. 20, n. 1, p. 45–57, 2001.
- ZHAO, W. et al. Trajectories of the hippocampal subfields atrophy in the alzheimer’s disease: A structural imaging study. **Frontiers in Neuroinformatics**, v. 13, n. March, p. 1–9, 2019. ISSN 16625196.

ZHENG, F. et al. The volume of hippocampal subfields in relation to decline of memory recall across the adult lifespan. **Frontiers in Aging Neuroscience**, v. 10, n. OCT, p. 1–10, 2018. ISSN 16634365.

Apêndices

Apêndice A

Informações Complementares

A.1 Ressonância Magnética

O imageamento por ressonância magnética (RM) é um método não-invasivo e não-ionizante de diagnóstico por imagem que possibilita a formação de imagens estruturais e funcionais de diferentes partes do corpo humano. A técnica de RM é capaz de diferenciar tecidos ao utilizar os átomos de hidrogênio presentes no corpo, sendo este escolhido por ser o átomo mais abundante no corpo humano e por possuir o maior momento magnético, o que reflete em uma maior sensibilidade ao IRM (MAZZOLA, 2009).

Podemos interpretar os átomos de hidrogênio como pequenos ímãs sujeitos a ação de campos magnéticos, devido às suas propriedades de momento magnético, e de *spin* ou momento angular, que descrevem o movimento de giro do próton em torno de seu próprio eixo enquanto produz um pequeno campo magnético. As partículas de hidrogênio presentes no corpo não possuem uma orientação espacial definida, logo, ao serem submetidas a um forte campo magnético de um *scanner* de RM, que usualmente opera em 1,5T ou 3T, elas irão se orientar (paralelamente ou antiparalelamente) de acordo com a direção do campo aplicado. A Figura 16 ilustra o processo de reorientação dos *spins* dos átomos de hidrogênio quando submetidos a um campo magnético externo.

Inicialmente um campo magnético é aplicado ao corpo para que os *spins* dos prótons fiquem alinhados a ele, em seguida, um segundo campo em pulsos de radiofrequência é aplicado perpendicular ao primeiro campo, modificando a orientação dos *spins* para o novo campo. Ao remover o segundo campo magnético, os átomos passam por um processo de relaxação para se alinharem novamente ao primeiro campo. Essa relaxação da orientação dos *spins* é causada pelas trocas de energia entre *spins* e sua vizinhança. A partir desses dois tipos de trocas de energia, foram definidas duas constantes de tempo de caracterização

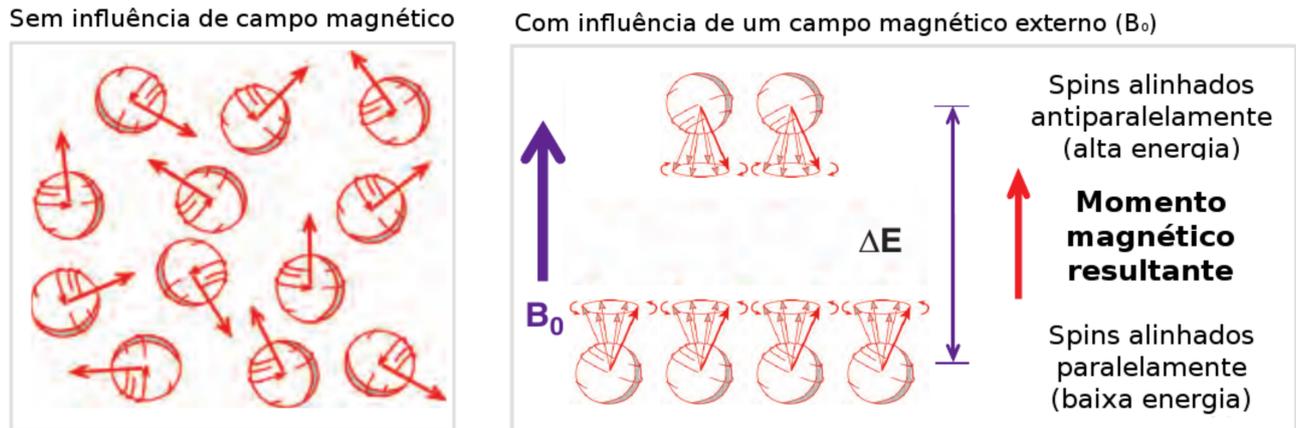


Figura 16 – Ilustração do alinhamento de prótons paralelamente e antiparalelamente à um campo magnético B_0 , extraído e adaptado de (BUSHBERG; BOONE, 2011).

do processo de relaxação, o T1 e T2. A constante T1 está relacionada ao tempo de retorno da magnetização para o eixo longitudinal e é influenciada pelos *spins* com a vizinhança, e já a constante T2 refere-se à redução da magnetização no plano transversal e é influenciada pela interação *spin-spin* (MAZZOLA, 2009). A Figura 17 apresenta uma ilustração do processo de relaxamento dos spins após a remoção do segundo campo magnético, logo, ocorre o retorno a orientação com o primeiro campo.

As imagens ponderadas em T1 possuem um melhor delineamento da anatomia das estruturas, enquanto que as imagens ponderadas em T2 possuem um maior contraste com lesões cerebrais e processos inflamatórios. No diagnóstico da doença de Alzheimer são utilizadas imagens de RM ponderadas em T1, devido a doença ser primariamente definida por alterações estruturais no cérebro.

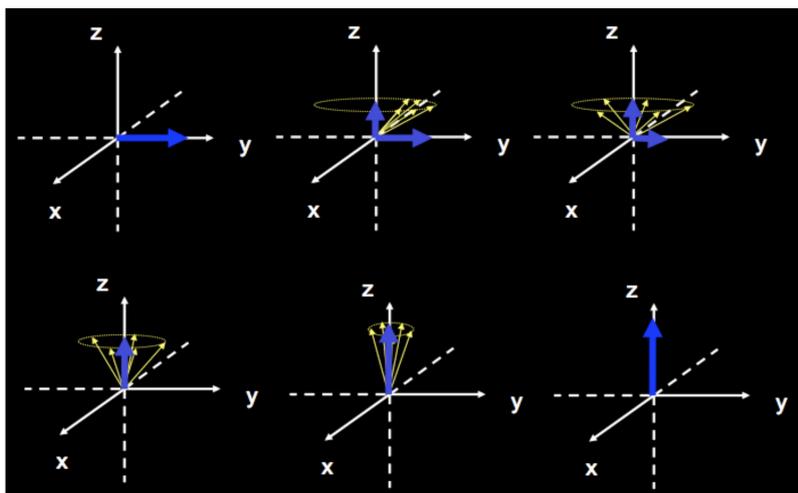


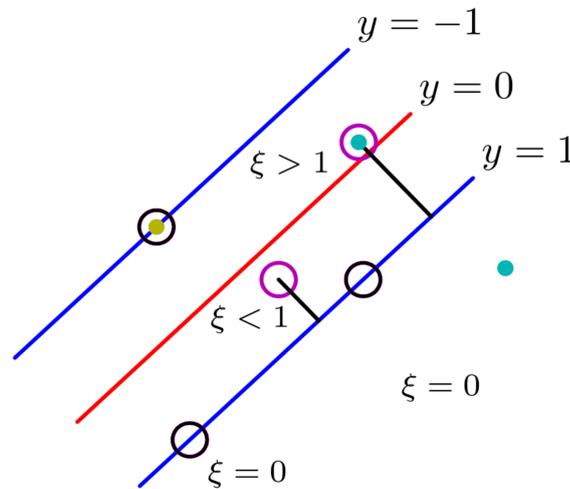
Figura 17 – Representação do processo de relaxamento dos prótons após a remoção do segundo campo magnético, extraído de (MAZZOLA, 2009).

A imagem de RM final é gerada por fatias do volume (corpo) de acordo com o método (a depender da fabricante do *scanner*) de aplicação do princípio de funcionamento descrito

nos parágrafos anteriores. Assim, cada imagem de RM é composta por voxels, a menor unidade da imagem de RM. Usualmente cada voxel representa $1,0\text{mm}^3$ do volume de aquisição, sendo que este valor depende do *scanner* utilizado. Os voxels possuem um valor de intensidade de cinza que variam de acordo com a densidade de prótons em cada tipo de tecido, sendo diretamente proporcional a intensidade da resposta do sinal de radiofrequência na região. (HANSON, 2009)

A.2 Classificador SVM

O SVM é um algoritmo de aprendizado de máquina, determinístico e com aprendizado supervisionado, usado para solucionar problemas de classificação linearmente separáveis ou não linearmente separáveis. O SVM linear busca encontrar um hiperplano linear que separe as amostras de treinamento pela maior distância do hiperplano para as amostras, com auxílio de dois hiperplanos que servem de margem (LORENA; CARVALHO, 2007). Para o caso não linear do SVM, busca-se uma curva que separe as amostras, sendo especialmente útil para dados não lineares, como é o caso dos grafos (KRIEGE; JOHANSSON; MORRIS, 2020b).



Fonte: Bishop (2006)

Figura 18 – Ilustração do funcionamento dos hiperplanos no SVM, com amostras em diferentes valores de ξ .

Dado um conjunto de vetores de treino $\mathbf{x}_i \in \mathbb{R}^n$, $i = 1, \dots, l$ em duas classes, e um vetor $\mathbf{y} \in \mathbb{R}^l$ tal que $\mathbf{y}_i \in \{-1, 1\}$ é o vetor de rótulos de cada amostra \mathbf{x}_i , o SVM não linear é obtido pela solução do problema de otimização primal (CORTES; VAPNIK, 1995),

$$\begin{aligned} \min_{w,b,\xi} \quad & \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{i=1}^l \xi_i \\ \text{sujeito à} \quad & y_i (\mathbf{w}^T \phi(\mathbf{x}_i) + b) \geq 1 - \xi_i, \\ & \xi_i \geq 0, i = 1, \dots, l \end{aligned} \quad (49)$$

em que \mathbf{w} é o vetor normal ao hiperplano, $\phi(\mathbf{x}_i)$ mapeia a amostra \mathbf{x}_i para um espaço de dimensão superior, $C > 0$ é o parâmetro de regularização que impõe um peso à minimização dos erros no conjunto de treinamento e ξ é chamada de variável de folga. A variável de folga ξ indica a possibilidade de amostras residirem entre a margem e o hiperplano separador, sendo considerado um erro quando $\xi > 1$. Logo, ao minimizar o termo $C \sum_{i=1}^l \xi_i$, minimiza-se os erros marginais do treinamento. Assim, a Equação 49 pode ser solucionada em sua versão dual como

$$\begin{aligned} \min_{\alpha} \quad & \frac{1}{2} \alpha^T Q \alpha - \mathbf{e}^T \alpha \\ \text{sujeito à} \quad & \mathbf{y}^T \alpha = 0, \\ & 0 \leq \alpha_i \leq C, i = 1, \dots, l \end{aligned} \quad (50)$$

para $\mathbf{e} = [1, \dots, 1]^T$, com α sendo os multiplicadores de Lagrange (utilizado para solução), Q é uma matriz $l \times l$ positiva e semidefinida, tal que $Q_{ij} = y_i y_j K(\mathbf{x}_i, \mathbf{x}_j)$ e $K(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j)$ é a função de *kernel*. Solucionando ambas equações 49 e 50 pelo relacionamento primal-dual, um \mathbf{w} ótimo, satisfaz

$$\mathbf{w} = \sum_{i=1}^l y_i \alpha_i \phi(\mathbf{x}_i), \quad (51)$$

e finalmente a função de decisão do SVM não-linear é dada por,

$$\text{sgn}(\mathbf{w}^T \phi(\mathbf{x}) + b) = \text{sgn}\left(\sum_{i=1}^l y_i \alpha_i K(\mathbf{x}_i, \mathbf{x}) + b\right), \quad (52)$$

Para o uso com um *kernel* pré computado, como o caso do *Graph Kernel*, o termo $K(\mathbf{x}_i, \mathbf{x})$ é substituído pela consulta à matriz de Gram definida na Equação 36, resultando em

$$\sum_{i=1}^l y_i \alpha_i G(\mathbf{x}_i, \mathbf{x}) + b. \quad (53)$$