



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS DA COMPUTAÇÃO

Rodrigo Gomes de Souza

Seleção automática de fatias em volumes de imagens tomográficas: Estudo de caso
no apoio ao diagnóstico da Doença de Alzheimer

Recife

2022

Rodrigo Gomes de Souza

Seleção automática de fatias em volumes de imagens tomográficas: Estudo de caso
no apoio ao diagnóstico da Doença de Alzheimer

Trabalho submetido ao Programa de Pós-Graduação em Ciência da Computação do Centro de Informática da Universidade Federal de Pernambuco como requisito parcial para obtenção do grau de Doutor em Ciência da Computação.

Área de Concentração: Inteligência Computacional

Orientador (a): Manoel Eusébio de Lima

Coorientador (a): Wellington Pinheiro dos Santos

Recife

2022

Catálogo na fonte
Bibliotecária Nataly Soares Leite Moro, CRB4-1722

S729s Souza, Rodrigo Gomes de
Seleção automática de fatias em volumes de imagens tomográficas: estudo de caso no apoio ao diagnóstico da Doença de Alzheimer / Rodrigo Gomes de Souza. – 2022.
117 f.: il., fig., tab.

Orientador: Manoel Eusébio de Lima.
Tese (Doutorado) – Universidade Federal de Pernambuco. CIn, Ciência da Computação, Recife, 2022.
Inclui referências.

1. Engenharia da Computação. 2. Seleção automática de fatias. 3. Imagens de ressonância magnética. 4. Computação evolucionária. 5. Computação bioinspirada. I. Lima, Manoel Eusébio de (orientador). II. Título

621.39 CDD (23. ed.) UFPE - CCEN 2022 – 203

Rodrigo Gomes de Souza

**“Seleção automática de fatias em volumes de imagens tomográficas:
Estudo de caso no apoio ao diagnóstico da Doença de Alzheimer”**

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutor em Ciência da Computação. Área de Concentração: Inteligência Computacional

Aprovado em: 30/06/2022.

Orientador: Prof. Dr. Manoel Eusebio de Lima

BANCA EXAMINADORA

Prof. Dr. Abel Guilhermino da Silva Filho
Centro de Informática/UFPE

Profa. Dra. Giselle Machado Magalhães Moreno
Instituto de Astronomia, Geofísica e Ciências Atmosféricas/USP

Prof. Dr. Marcelo Cairrão Araujo Rodrigues
Departamento de Fisiologia e Farmacologia/UFPE

Prof. Dr. Ricardo Emmanuel de Souza
Departamento de Engenharia Biomédica/UFPE

Prof. Dr. Sidney Marlon Lopes de Lima
Departamento de Eletrônica e Sistemas / UFPE

Dedico este trabalho a minha esposa e filhas e também aos meus pais pela fé que todos sempre depositaram em mim.

AGRADECIMENTOS

Primeiramente, agradeço a Deus por ter colocado em meu caminho tantas pessoas iluminadas e especiais - sem as quais a realização desse trabalho não seria possível. Também gostaria de agradecer toda minha família pelo amor e carinhos concedidos, em especial a minha esposa Caroline e as minhas filhas Esther e Júlia. Por fim, agradeço aos meus orientadores Manoel e Wellington, os quais foram fundamentais na conclusão desta tese.

A educação faz sentido porque as mulheres e homens aprendem que através da aprendizagem podem fazerem-se e refazerem-se, porque mulheres e homens são capazes de assumirem a responsabilidade sobre si mesmos como seres capazes de conhecerem.

(FREIRE, 2001)

RESUMO

Com o advento da Quarta Revolução Industrial, as tecnologias da Indústria 4.0 têm acelerado o processo de transformação digital dos sistemas nacionais de saúde, tanto públicos quanto complementares, incorporando a Inteligência Artificial, a Internet das Coisas, a Robótica e a Biotecnologia à prática clínica, terapêutica e de gestão de recursos em saúde. Com a pandemia de Covid-19, também se intensificou o uso de técnicas avançadas de diagnóstico por imagem. Apesar da sofisticação do diagnóstico por imagem por conta da possibilidade de gerar laudos rápidos e precisos, esses exames exigem um alto conhecimento especialista. Contudo, a dependência do conhecimento especialista humano pode acarretar em diagnósticos tardios e, por conseguinte, maus prognósticos. Esse problema é especialmente mais grave quando se consideram imagens volumétricas: estruturas associadas a determinadas doenças com bons prognósticos quando detectadas precocemente podem ser ignoradas até mesmo por especialistas com grande experiência, por razões diversas, incluindo aí a fadiga humana, dada a complexidade envolvida na análise humana de volumes. Devido ao potencial dos impactos à saúde pública e à economia de forma geral, a Doença de Alzheimer, DA, tornou-se um dos maiores desafios à ciência na última década. Contudo, grande parte dos estudos que propõem métodos de suporte ao diagnóstico da doença apresentam dificuldades para sua adoção na prática clínica. Este trabalho tem como objetivo geral construir uma metodologia para selecionar planos de visualização e cortes ou fatias em imagens biomédicas volumétricas, de forma que se otimize o diagnóstico por meio da transformação de um problema tridimensional em bidimensional. A seleção da fatia e dos planos de visualização é feita com o auxílio de um algoritmo evolucionário guiado por um classificador. Como estudo de caso, foi escolhido o problema do apoio ao diagnóstico da doença de Alzheimer. Utilizando volumes MRI da base ADNI-1, este estudo apresenta um modelo baseado em computação evolucionária e aprendizado de máquina capaz de identificar e utilizar o conjunto de fatias 2D mais adequado para maximizar cada uma das métricas relacionadas ao diagnóstico da doença de Alzheimer. Por meio de uma seleção ótima de fatias os melhores resultados para especificidade, sensibilidade e acurácia obtidos em cada classe foram respectivamente: normal (94,58%, 81,08%, 86,64%), déficit cognitivo leve (94,83%, 89,32%, 89,89%) e Alzheimer (92,99%, 92,06%, 90,61%) para mulheres, e normal (91,93%, 93,42%, 91,30%), déficit cognitivo leve (98,17%, 80,53%, 86,29%) e Alzheimer (95,86%, 90,91%, 94,31%) para homens. Assim, este trabalho apresenta um modelo para apoiar o diagnóstico de doença de Alzheimer e déficit cognitivo

leve, apresentando bom desempenho de classificação, considerando também as informações de gênero para produzir melhorias significativas no diagnóstico da doença de Alzheimer.

Palavras-chaves: seleção automática de fatias; imagens de ressonância magnética; computação evolucionária; computação bioinspirada; doença de alzheimer.

ABSTRACT

With the advent of the Fourth Industrial Revolution, Industry 4.0 technologies have accelerated the process of digital transformation of national health systems, both public and complementary, incorporating Artificial Intelligence, the Internet of Things, Robotics and Biotechnology into clinical practice, therapy and management of health resources. With the Covid-19 pandemic, the use of advanced imaging diagnostic techniques has also intensified. Despite the sophistication of diagnostic imaging due to the possibility of generating fast and accurate reports, these exams require high specialist knowledge. However, dependence on human expert knowledge can lead to late diagnoses and, therefore, poor prognosis. This problem is especially more serious when considering volumetric images: structures associated with certain diseases with good prognoses when detected early can be ignored even by specialists with great experience, for various reasons, including human fatigue, given the complexity involved in the analysis. volume human. Due to the potential impacts on public health and the economy in general, Alzheimer's Disease, AD, has become one of the greatest challenges to science in the last decade. However, most of the studies that propose methods to support the diagnosis of the disease present difficulties in their adoption in clinical practice. The general objective of this work is to build a methodology to select visualization planes and cuts or slices in volumetric biomedical images, in a way that optimizes the diagnosis by transforming a three-dimensional problem into a two-dimensional one. The selection of slice and view planes is done with the aid of an evolutionary algorithm guided by a classifier. As a case study, the problem of support for the diagnosis of Alzheimer's disease was chosen. Using MRI volumes from the ADNI-1 base, this study presents a model based on evolutionary computation and machine learning capable of identifying and using the most suitable set of 2D slices to maximize each of the metrics related to the diagnosis of Alzheimer's disease. Through an optimal selection of slices, the best results for specificity, sensitivity and accuracy obtained in each class were respectively: normal (94.58%, 81.08%, 86.64%), mild cognitive impairment (94.83 %, 89.32%, 89.89%) and Alzheimer's (92.99%, 92.06%, 90.61%) for women, and normal (91.93%, 93, 42%, 91.30%), mild cognitive impairment (98.17%, 80.53%, 86.29%) and Alzheimer's (95.86%, 90.91%, 94.31%) for men. Thus, this work presents a model to support the diagnosis of Alzheimer's disease and mild cognitive impairment, presenting good classification performance, also considering gender information to produce significant improvements in the diagnosis of Alzheimer's disease.

Keywords: automatic slice selection; magnetic resonance images; evolutionary computing; bioinspired computing; alzheimer's disease.

LISTA DE FIGURAS

Figura 1 – Modelo da progressão da Doença de Alzheimer	26
Figura 2 – Modelo Discriminante construído a partir de aprendizagem supervisionada.	30
Figura 3 – Interação entre agente e ambiente.	32
Figura 4 – Um Modelo construído por SVM de Margem Rígida	33
Figura 5 – Infinitos separadores lineares separam duas classes linearmente separáveis	34
Figura 6 – Cálculo da distância entre hiperplanos	35
Figura 7 – Classificação com restrições relaxadas	39
Figura 8 – Exemplo de convolução sem inversão de <i>kernel</i>	43
Figura 9 – Estágios que compõem uma camada típica de uma rede neural convolucional	45
Figura 10 – Função <i>Maxpooling</i>	46
Figura 11 – Proposta: Fluxograma de aprendizado de máquina aplicado para avaliar um agrupamento específico de slices.	59
Figura 12 – Uso: classificação de volume de acordo com nossa proposta.	60
Figura 13 – Exemplos de fatias de volumes ponderados em T1 de indivíduos do sexo masculino cognitivamente normais, para visualizações coronal (1, 2 e 3), sagital (4, 5 e 6) e axial (7, 8 e 9).	68
Figura 14 – Exemplos de fatias de volumes ponderados em T1 para indivíduos do sexo masculino com comprometimento cognitivo leve, para visualizações coronal (1, 2 e 3), sagital (4, 5 e 6) e axial (7, 8 e 9).	69
Figura 15 – Exemplos de fatias de volume ponderados em T1 para indivíduos do sexo masculino com doença de Alzheimer, para visualização coronal (1, 2 e 3), sagital (4, 5 e 6) e axial (7, 8 e 9).	70
Figura 16 – Exemplos de fatias de volume ponderados em T1 para indivíduos do sexo feminino cognitivamente normais, para visualizações coronal (1, 2 e 3), sagital (4, 5 e 6) e axial (7, 8 e 9).	71
Figura 17 – Exemplos de fatias de volume ponderados em T1 para mulheres com comprometimento cognitivo leve, para vistas coronal (1, 2 e 3), sagital (4, 5 e 6) e axial (7, 8 e 9).	72

Figura 18 – Exemplos de fatias de volume ponderados em T1 para mulheres com doença de Alzheimer, para vistas coronal (1, 2 e 3), sagital (4, 5 e 6) e axial (7, 8 e 9).	73
Figura 19 – Exemplos de fatias de volumes ponderados em T1 após BET para indivíduos do sexo masculino cognitivamente normais, para visualizações coronal (1, 2 e 3), sagital (4, 5 e 6) e axial (7, 8 e 9).	77
Figura 20 – Exemplos de fatias de volumes ponderados em T1 após BET para indivíduos do sexo masculino com comprometimento cognitivo leve, para visualizações coronal (1, 2 e 3), sagital (4, 5 e 6) e axial (7, 8 e 9)	78
Figura 21 – Exemplos de fatias de volumes ponderados em T1 após BET para indivíduos do sexo masculino com doença de Alzheimer, para vistas coronal (1, 2 e 3), sagital (4, 5 e 6) e axial (7, 8 e 9).	79
Figura 22 – Exemplos de fatias de volumes ponderados em T1 após BET para mulheres cognitivamente normais, para vistas coronal (1, 2 e 3), sagital (4, 5 e 6) e axial (7, 8 e 9).	80
Figura 23 – Exemplos de fatias de volumes ponderados em T1 após BET para indivíduos do sexo feminino com comprometimento cognitivo leve, para visualizações coronal (1, 2 e 3), sagital (4, 5 e 6) e axial (7, 8 e 9).	81
Figura 24 – Exemplos de fatias de volumes ponderados em T1 após BET para mulheres com doença de Alzheimer, para vistas coronal (1, 2 e 3), sagital (4, 5 e 6) e axial (7, 8 e 9).	82
Figura 25 – Histogramas de distribuição de fatias para participantes do sexo feminino (1), masculino (2) e de ambos os sexos (3), considerando o pré-processamento da segmentação BET.	87
Figura 26 – Histogramas de distribuição de fatias para participantes do sexo feminino (1), masculino (2) e de ambos os sexos (3), sem segmentação BET.	88

LISTA DE TABELAS

Tabela 1 – Parâmetros definidos para os classificadores empregados	61
Tabela 2 – Resultados agrupados por gênero, usando o pré-processamento BET	85
Tabela 3 – Resultados agrupados por gênero, sem uso de BET	86
Tabela 4 – Cortes mais frequentes e frequência absoluta e relativa, segundo sexo (F para feminino, M para masculino e ambos) e plano corporal com pré-processamento BET	89
Tabela 5 – Cortes mais frequentes em frequência absoluta e relativa, segundo sexo (F para feminino, M para masculino e ambos) e plano corporal sem pré-processamento BET	90
Tabela 6 – Resultados do teste de normalidade Kolmogorov-Smirnov das distribuições de fatias de acordo com o sexo, com e sem segmentação por BET.	90
Tabela 7 – Resultados da hipótese de Kruskal-Wallis para k amostras independentes	91
Tabela 8 – Resultados dos testes de hipóteses bicaudais de Mann-Whitney Wilcoxon.	91

LISTA DE ABREVIATURAS E SIGLAS

AAL	Automated Anatomical Labelling
AD	Alzheimer's Disease
ADAS-Cog13	Alzheimer's Disease Assessment Scale-Cognitive
ADNI	Alzheimer's Disease Neuroimaging Initiative
ANART	American National Adult Reading Test
ApoE	Apolipoproteína E
APP	Amyloid Precursor Protein
AUC	Área Sobre a Curva
BET	Brain Extraction Tool
BHR	Brain Health Registry
c-MCI	MCI que converge para Alzheimer
CART	Árvores de Decisão
CBB	Cogstate Brief Battery
CCI	Índice de Mudança Cognitiva
CCL	Comprometimento Cognitivo Leve
CDR	Clinical Dementia Rating Scale
CHF	Circular Harmonic Functions
CN	Cognitivamente Normal
CNN	Convolutional Neural Networks
CSF	Líquido Cefalorraquidiano (Líquido Espinhal Cerebral)
DA	Doença de Alzheimer
DSC	Índice de Sørensen–Dice

ECog	Medição da Cognição Cotidiana
f-MRI	Imageamento por Ressonância Magnética Funcional
FCI-SF	Financial Capacity Instrument – Short Form
FDG-PET	Fluoro-Deoxi-Glicose PET
GDS	Geriatric Depression Scale
GM	Massa (Cerebral) Cinzenta
ICC	Coefficiente de Correlação Intraclasse
KNN	K-Nearest Neighbors
LDA	Análise Discriminante Linear
LEC	Córtex Entorrinal Lateral
LMBP	Levenberg–Marquardt (LM) back-propagation
LMCI	MCI Tardio
LOO	(Validação) LeaveOne-Out
LVQ	Learning Vector Quantization
MCI	Prejuízo Cognitivo Moderado ou Leve
MEC	Córtex Entorrinal Medial
ML	Aprendizado por Máquina ou Machine Learning
MINT	Multi-Lingual Naming Test
MMSE	Mini-Mental State Examinations
MNI	Montreal Neurological Institute
MoCA	Montreal Cognitive Assessment
MR	Magnetic Resonance
NB	Naive Bayes

NC	Normal Control
NPI	Inventário Neuropsiquiátrico
NPI-Q	Inventário Neuropsiquiátrico Rápido
OASIS	Open Access Series of Imaging Studies
OCR	Reconhecimento Óptico de Caracteres
PCA	Principal Component Analysis
PET	Positron Emission Tomography
PLS	Partial Least Squares
pMCI	Progressive MCI
PMZ	Pseudo Momentos de Zernick
PRNN,	Pattern Recognition Neural Network
RBF	Radial Basis Function
RBP	Resiliente Back Propagation
RF	Random Forest
RL	Regressão Logística
RM	Ressonância Magnética
ROI	Região de Interesse
s-MRI	Imageamento por Ressonância Magnética Estrutural
SCG	Scaled Conjugate Gradient
sMCI	MCI estável ou que não evolui
SMOTE	Synthetic Minority Oversampling Technique
SVM	Máquina de Vetor de Suporte
VIQ	Inteligência Verbal Pré-Mórbida

WM Massa (Cerebral) Branca

WMS-R Wechsler Memory Scale-Revised

SUMÁRIO

1	INTRODUÇÃO	20
1.1	MOTIVAÇÃO E JUSTIFICATIVA	20
1.2	OBJETIVOS	23
1.3	ORGANIZAÇÃO DA TESE	24
2	FUNDAMENTAÇÃO TEÓRICA	25
2.1	A DOENÇA DE ALZHEIMER	25
2.1.1	Modelagem da progressão da doença	25
2.1.2	Biomarcadores adquiridos por Imagens	27
2.2	APRENDIZADO DE MÁQUINA	28
2.2.1	Tipos de Aprendizado	29
2.2.1.1	<i>Aprendizado Supervisionado</i>	29
2.2.1.2	<i>Aprendizado Não Supervisionado</i>	31
2.2.1.3	<i>Aprendizado por Reforço</i>	31
2.3	MÁQUINAS DE VETOR DE SUPORTE	32
2.3.1	SVMs de Margem Rígida	33
2.3.2	SVMs de Margem Flexível	38
2.4	REDES NEURAIAS CONVOLUCIONAIS	41
2.4.1	A Operação de Convolução	41
2.4.2	Características das redes neurais convolucionais	43
2.4.3	Arquitetura de uma rede convolucional	44
2.5	CONSIDERAÇÕES	46
3	TRABALHOS RELACIONADOS	47
3.1	ESTADO DA ARTE	47
3.2	CONSIDERAÇÕES	53
4	PROPOSTA, MATERIAIS E MÉTODOS	55
4.1	REPRESENTAÇÃO DE IMAGENS	55
4.2	MODELAGEM DA OTIMIZAÇÃO	56
4.3	MODELAGEM DOS CANDIDATOS À SOLUÇÃO	56
4.4	FUNÇÃO OBJETIVO	57
4.5	PIPELINE DE APRENDIZADO DE MÁQUINA	57

4.5.1	Formato de instâncias de treinamento e teste	57
4.5.2	Seleção de características e redução de dimensionalidade	58
4.5.3	Balanceamento de classes	58
4.5.4	Validação do modelo	58
4.6	PROJETO DE EXPERIMENTOS	60
4.7	BASE DE DADOS	63
4.8	PRÉ-PROCESSAMENTO: EXTRAÇÃO CEREBRAL AUTOMATIZADA . .	74
4.9	BALANCEAMENTO DE CLASSES MINORITÁRIA POR MEIO DO MÉ- TODO SMOTE	83
4.10	ARRANJO EXPERIMENTAL	84
5	RESULTADOS E DISCUSSÃO	85
5.1	RESULTADOS EXPERIMENTAIS	85
5.2	DISCUSSÃO	92
6	CONCLUSÃO	98
6.1	CONCLUSÕES GERAIS	98
6.2	DIFICULDADES APRESENTADAS	99
6.3	CONTRIBUIÇÕES E TRABALHOS FUTUROS	100
	REFERÊNCIAS	102

1 INTRODUÇÃO

Este capítulo apresenta a introdução à Tese de Doutorado, sendo apresentados o contexto, a justificativa e a motivação para a realização da pesquisa relatada, bem como o objetivo geral e os objetivos específicos ou metas.

1.1 MOTIVAÇÃO E JUSTIFICATIVA

A Transformação Digital da Saúde consiste na mudança de paradigma tanto nos serviços de promoção da saúde e nas práticas curativas e terapêuticas quanto na gestão desses serviços e recursos associados em direção ao uso mais intensivo das Tecnologias da Informação e da Comunicação, com a transformação de práticas de diagnóstico, terapia e gestão de recursos de saúde e a consequente criação de novas práticas (BALTAXE et al., 2019; ADLER-MILSTEIN, 2021; TSENG et al., 2018; DZAU; BALATBAT, 2019). Com o advento da Quarta Revolução Industrial, as tecnologias da Indústria 4.0 têm acelerado o processo de transformação digital dos sistemas nacionais de saúde, tanto públicos quanto complementares, incorporando a Inteligência Artificial, a Internet das Coisas, a Robótica e a Biotecnologia à prática clínica, terapêutica e de gestão de recursos em saúde (JAYARAMAN et al., 2020; CHUTE; FRENCH, 2019).

A pandemia de Covid-19, iniciada na China em dezembro de 2019, acelerou o processo de transformação digital da saúde no mundo, fazendo com que surgissem diversas aplicações nos campos da telessaúde, da telemedicina, e da inteligência artificial, em especial no apoio ao diagnóstico. Para dar suporte legal a essas diversas tecnologias, países que estavam atrasados na regulamentação da telemedicina e da saúde regulamentaram o seu uso e já estudam formas de incluir as ferramentas de inteligência artificial (STEIN; NASLUND; BANTJES, 2022; PETRACCA et al., 2020).

Com a pandemia de Covid-19, também se intensificou o uso de técnicas avançadas de diagnóstico por imagem, como a tomografia computadorizada por raios-x, utilizada para diagnosticar a Covid-19 e acompanhar o progresso do tratamento da doença (GOMES et al., 2020b; LOW et al., 2021). Diversos grupos de pesquisa, grandes empresas e *healthtechs* em todo o mundo buscaram construir ferramentas para apoio ao diagnóstico por imagem, agregando valor aos sistemas de armazenamento de imagens clínicas existentes por meio da introdução de módulos inteligentes (GOMES et al., 2020b; LOW et al., 2021; SANTANA et al., 2018; RODRIGUES

et al., 2019; SANTANA et al., 2020; SILVA et al., 2021; SANTANA et al., 2021).

Contudo, apesar da sofisticação do diagnóstico por imagem por conta da possibilidade de gerar laudos rápidos e precisos, esses exames exigem um alto conhecimento especialista. No entanto, o contexto de pandemia mostrou que, em situações emergenciais, por exemplo, a dependência do conhecimento especialista humano pode acarretar em diagnósticos tardios e, por conseguinte, maus prognósticos. Esse problema é especialmente mais grave quando se consideram imagens volumétricas: estruturas associadas a determinadas doenças com bons prognósticos quando detectadas precocemente podem ser ignoradas até mesmo por especialistas com grande experiência, por razões diversas, incluindo aí a fadiga humana, dada a complexidade envolvida na análise humana de volumes (BOGNER, 2018; RAZZAK; NAZ; ZAIB, 2018; LEE et al., 2018).

Uma possível solução para este problema pode passar pela transformação do problema de análise 3D em um problema 2D, mas como selecionar as melhores fatias ou cortes em um volume de imagens, e como escolher os melhores planos de visualização, de forma a otimizar o processo de classificação e, por conseguinte, o processo de diagnóstico? Esta é a pergunta de pesquisa deste trabalho.

A hipótese de pesquisa desta tese consiste na seguinte: algoritmos evolucionários podem ser utilizados na seleção de planos de visualização e de fatias ou cortes que otimizem o apoio ao diagnóstico de doenças a partir de imagens biomédicas volumétricas, transformando um problema de análise tridimensional em um problema bidimensional e otimizando as métricas de qualidade de classificação.

A demência, frente ao aumento da longevidade populacional, tornou-se a ameaça mais temida pela sociedade, representando um problema de saúde pública urgente e prioritário. Até por volta de 1968, a Doença de Alzheimer (DA) era considerada uma forma rara de demência pré-senil, diferentemente da demência senil já era considerada bastante comum (CASELLI et al., 2006). A partir de trabalhos publicados na época, foi possível identificar correlações entre a demência senil e os traços neuropatológicos marcantes da DA, apontando semelhanças nas condições clínicas entre as duas patologias (BLESSED; TOMLINSON; ROTH, 1968). Por isso, em 1976, decidiu-se unir os dois conceitos em um só, com o termo demência senil do tipo Alzheimer. A DA então passou de uma posição considerada rara para o outro extremo, na qual se estabeleceu praticamente em uma epidemia hoje, tomando para si quase todos os diagnósticos que se enquadram em uma condição progressiva de degeneração do tecido cerebral (HODGES, 2006). O conceito de DA evoluiu e continua evoluindo constantemente,

pois novas descobertas no campo da neurociência são frequentes, direcionando e modificando a perspectiva do diagnóstico e dos processos terapêuticos. No início do século XXI, e com o aumento da esperança de vida, tem havido um interesse crescente em diagnosticar a DA com a maior precisão possível, uma vez que a incidência de demências, e a DA em particular, tem vindo a aumentar progressivamente, causando não só doenças individuais e problemas de saúde familiar, mas também sociais e econômicos.

A doença de Alzheimer faz parte do pequeno grupo de doenças cujo diagnóstico preciso é confirmado através do estudo histopatológico post-mortem obtido por biópsias ou necropsia cerebral (HAMDAN, 2008). No entanto, o diagnóstico baseado em evidências científicas realmente cria a oportunidade para que medidas terapêuticas adequadas, sejam com medicamentos ou não, sejam aplicadas antecipadamente. Além disso, permite que familiares e pacientes programem seu tempo de forma mais eficiente, não apenas para que se preparem para lidar com a situação que se aproxima, mas para que busquem os recursos que serão necessários. Neste caso, destacamos o uso de Imagens de Ressonância Magnética (RM), mais indicadas para visualizar alterações anatômicas na estrutura cerebral.

A RM é adquirida por método não invasivo, sem uso de radiação, o que tem grande vantagem em relação a outros exames de imagem com tecnologia mais simples, como a radiografia, por exemplo, devido ao seu contraste, embora persista como uma técnica consideravelmente cara. A qualidade da imagem de ressonância magnética permite a diferenciação do tecido e a identificação estrutural para ajudar a localizar patologias. Sua tecnologia baseia-se na magnetização natural dos átomos do corpo, principalmente do hidrogênio. Este exame gera imagens volumétricas, nos três planos do corpo (axial, coronal e sagital). Assim é possível investigar áreas específicas ligadas ao traço patológico neurodegenerativo particularmente ligado à doença de Alzheimer.

Os crescentes impactos da doença, especialmente em países mais desenvolvidos, onde a média de idade da população é mais alta, incentivou o estabelecimento de iniciativas globais para o desenvolvimento de biomarcadores como ADNI, Alzheimer's Disease Neuroimaging Initiative (RAMÍREZ et al., 2018; ALAM; KWON; INITIATIVE, 2017; JR et al., 2008; PETERSEN et al., 2010; MUELLER et al., 2005), e AIBL, Australian Imaging, Biomarker & Lifestyle Flagship Study of Aging (CARDOSO et al., 2017; ELLIS et al., 2014; ROWE et al., 2010; ELLIS et al., 2009). Esses trabalhos forneceram uma grande quantidade de dados longitudinais, de natureza clínica e demográfica, que poderiam ser utilizados por estudos baseados em aprendizado de máquina para desenvolver modelos multiclasse de diagnóstico precoce mais precisos.

O aprendizado de máquina é usado principalmente para tarefas de reconhecimento de padrões. Esses padrões devem estar localizados em um determinado conjunto de dados, seja em imagens, textos ou dados numéricos por exemplo. Um algoritmo para ser considerado como aprendizado deve ser capaz de: (1) quando exposto a um determinado conjunto de dados, detectar e aprender padrões, e (2) quando exposto a novos dados, e já com o conhecimento obtido anteriormente do conjunto onde os padrões foram identificados, classificar novas informações. Em resumo, os algoritmos realizam o processo de treinamento nos bancos de dados para identificar padrões e são capazes de aprender com eles, então reconhecer esses mesmos padrões em uma nova coleta de dados (GOMES et al., 2020c; SILVA et al., 2020b; GOMES et al., 2020a; LIMA; SILVA-FILHO; SANTOS, 2016a; COMMOWICK et al., 2018; CORDEIRO; SANTOS; SILVA-FILHO, 2016; SANTANA et al., 2018; GOMES et al., 2020b).

Dada a importância desse tema e seguindo a tendência de pesquisas que seguem a linha da inteligência artificial (SANTOS et al., 2009b; SANTOS et al., 2009a; SANTOS et al., 2009c; SANTOS et al., 2008; ISLAM; ZHANG, 2017; SILVA et al., 2020b), neste trabalho é proposta a análise automatizada de volumes de ressonância magnética da cabeça usando algoritmos de aprendizado de máquina pela seleção automática das fatias de MR relevantes capazes de otimizar os resultados da classificação usando a pesquisa evolutiva. Os modelos de classificação foram definidos de acordo com o sexo, uma vez que as demências em geral, especialmente a doença de Alzheimer, tendem a afetar indivíduos do sexo masculino e feminino de forma diferente, com efeitos mais graves sobre as mulheres (LAWS; IRVINE; GALE, 2016; HENDERSON; BUCKWALTER, 1994; OTT et al., 1996; EIKELBOOM et al., 2022; MARTINKOVA et al., 2021).

Desta forma, a hipótese de pesquisa deste trabalho será validada considerando o apoio ao diagnóstico da doença de Alzheimer por meio de neuroimagens volumétricas de ressonância magnética ponderadas em T1 como estudo de caso.

1.2 OBJETIVOS

Este trabalho tem como objetivo geral construir uma metodologia para selecionar planos de visualização e cortes ou fatias em imagens biomédicas volumétricas, de forma que se otimize o diagnóstico por meio da transformação de um problema tridimensional em bidimensional. A seleção da fatia e dos planos de visualização é feita com o auxílio de um algoritmo evolucionário guiado por um classificador. Como estudo de caso, foi escolhido o problema do apoio ao diagnóstico da doença de Alzheimer a partir de volumes de imagens do cérebro adquiridas

com ressonância magnética ponderada em T1. As imagens foram cedidas pelo Projeto ADNI.

Como objetivos específicos, têm-se:

1. Construir uma metodologia para representar volumes de neuroimagens volumétricas de ressonância magnética na forma de matrizes de atributos, sendo um vetor de atributos por fatia;
2. Construir uma arquitetura para seleção automática de fatias em neuroimagens volumétricas de ressonância magnética por computação evolucionária;
3. Construir uma arquitetura para seleção automática de planos de visualização e de fatias para apoio ao diagnóstico da doença de Alzheimer;
4. Investigar as melhores abordagens de classificação em função do sexo dos participantes.

1.3 ORGANIZAÇÃO DA TESE

Esta Tese está organizada da forma que segue:

No Capítulo 2 é apresentada a fundamentação teórica, sendo apresentada a definição da doença de Alzheimer, para fundamentação do estudo de caso desta tese, e uma breve revisão de princípios, definições e métodos de aprendizado de máquina.

No Capítulo 3 é apresentado o estado da arte das técnicas e metodologias para apoio ao diagnóstico da doença de Alzheimer por meio de aprendizado de máquina, bem como as bases de dados, com foco especial nas bases de imagens de ressonância magnética, tal como a abordagem do estudo de caso desta pesquisa.

A proposta deste trabalho é apresentada no Capítulo 4 por meio de descrições metodológicas detalhadas e diagramas de blocos. Também são apresentadas as limitações do estudo e as bases de dados utilizadas.

Os resultados experimentais são apresentados no Capítulo 5, para definição das melhores fatias e dos melhores planos de visualização de neuroimagens volumétricas de ressonância magnética ponderadas em T1 que otimizam o apoio ao diagnóstico da doença de Alzheimer.

Por fim, as dificuldades apresentadas até o momento, as conclusões parciais, contribuições e trabalhos futuros são apresentados no Capítulo 6.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo é apresentada a fundamentação teórica da doença de Alzheimer, o estudo de caso deste trabalho, e de princípios de aprendizado de máquina.

2.1 A DOENÇA DE ALZHEIMER

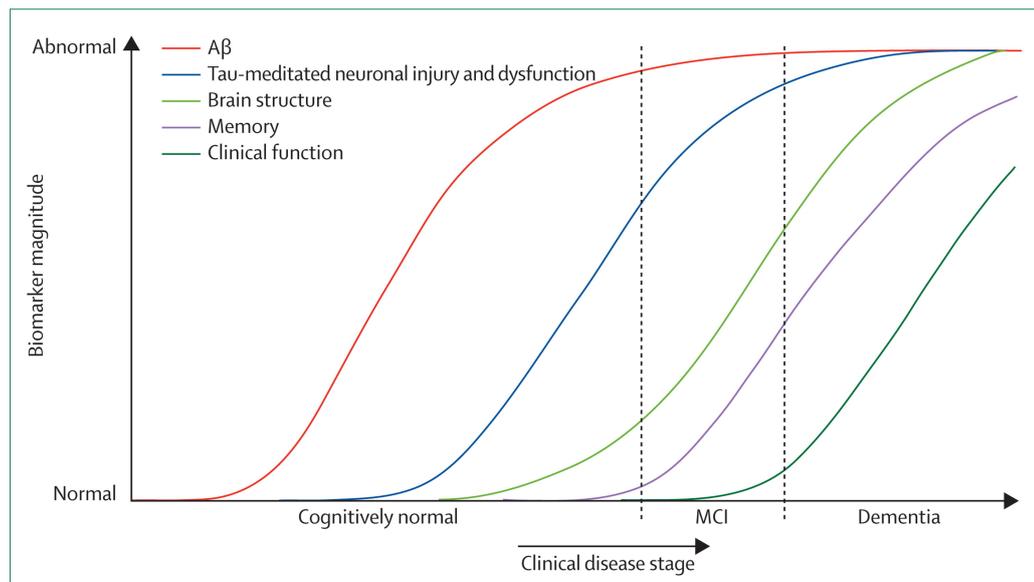
Ao longo dos últimos anos, diversos estudos dedicaram-se a identificar biomarcadores (indicadores mensuráveis da severidade ou do estado de uma doença) relacionados como os estágios da doença de Alzheimer (YOUNG et al., 2014). Além disso, tais marcadores têm sido utilizados, além de outros propósitos, para também estudar outros fatores importantes a respeito do espectro da doença para compreender quais regiões são primeiro afetadas pela neurodegeneração (SCHMITZ et al., 2016). O modelo mais amplamente aceito é conhecido como Hipótese Beta-Amilóide e descrito na próxima seção.

2.1.1 Modelagem da progressão da doença

Na Doença de Alzheimer, células que processam, armazenam e recuperação informação degeneram e morrem e as causas relacionadas com esse processo destrutivo ainda são desconhecidas. Porém, uma das teorias mais aceitas, conhecida como Hipótese da Beta-Amilóide, sugere como principal causa uma proteína chamada beta-amilóide a qual acumula-se no cérebro destruindo a comunicação entre as células e eventualmente levando-as à morte. Esta proteína é um pequeno pedaço de outra maior: a proteína precursora amilóide (*amyloid precursor protein, APP*). Em sua forma completa, a APP estende-se do interior até o exterior da célula nervosa, passando através da membrana de gordura que a envolve. Quando a APP está ativada para realizar sua função normal, ela é cortada por outras proteínas em pedaços menores que permanecem dentro e fora das células. Existem diversas formas que a APP pode ser dividida e, sob algumas circunstâncias, uma das partes produzidas é a beta-amilóide. Contudo, a beta-amilóide é um composto quimicamente mais pegajoso que os demais fragmentos resultantes da divisão da APP. Assim, ela acaba acumulando-se em etapas formando placas beta-amilóide e emaranhados neurofibrilares compostos de fibras tau-amilóide (HARDY, 2006; ASSOCIATION, 2017).

Jack-Jr et al. (2010) propuseram um modelo de progressão da doença que é empiricamente suportado por evidências apresentadas em diversos trabalhos (OKONKWO et al., 2010; MARSHALL et al., 2011; TOSUN et al., 2010; SHAW et al., 2011; SCHUFF et al., 2009). Este modelo, apresentado na Figura 1, é baseado em uma série de eventos patológicos que levam a prejuízos cognitivos e, depois, à demência.

Figura 1 – Modelo da progressão da Doença de Alzheimer



Fonte: Jack-Jr et al. (2010)

Com isso, a progressão da doença é definida em três estágios principais que são identificados da seguinte forma:

- Cognitivamente Normal (CN): apresenta acúmulo anormal de fibrilas de beta-amilóide, evidenciado pela redução nos níveis de solubilidade da beta-amilóide 42 no líquido cefalorraquidiano (Cerebrospinal Fluid, CSF) com aumento de retenção do indicador C-PiB no córtex evidenciado na Tomografia por Emissão de Pósitrons (PET);
- Prejuízo Cognitivo Moderado ou Leve (MCI): depois de um tempo, os danos neurais começam a ocorrer. Ocorre um aumento no nível da proteína tau no CSF. A disfunção sináptica começa a ocorrer, resultando na redução da fluoro-deoxiglicose observada pelo imageamento por PET (FDG-PET);
- Demência: A degeneração neural prossegue e a atrofia em certas áreas do cérebro, típicas de AD, são detectadas por meio de imageamento por ressonância magnética (MRI).

Portanto, o MCI é definido como o estágio no qual o paciente possui uma deficiência cognitiva maior que o esperado para sua idade, mas sem afetar suas tarefas diárias, e que ocorre antes do estado conhecido por “Demência por Doença de Alzheimer”, no qual a habilidade em executar tarefas diárias do cotidiano é, de fato, afetada. Segundo Zhang et al. (2014), o MCI é uma condição mais complexa e bastante heterogênea. Os autores, em sua revisão, afirmam que existem diversas definições para o MCI (PETERSEN et al., 1999; PETERSEN, 2004; WINBLAD et al., 2004; MATTHEWS et al., 2008). Por fim, existem quatro desfechos possíveis para o MCI: (i) progressão para a demência AD, (ii) progressão para outro tipo de demência, (iii) manutenção estável do MCI (ou *stable-MCI*, *s-MCI*) ou (iv) recuperação da capacidade cognitiva.

Além da teoria da beta-amilóide, existem outros estudos que sugerem a utilização de outros biomarcadores para identificar as fases da doença. Em sua revisão, Weiner et al. (2013) citam duas instituições norte americanas (Instituto Nacional do Envelhecimento e a Associação Alzheimer) como responsáveis por desenvolver um critério para identificar casos, a partir de biomarcadores, onde o prejuízo cognitivo moderado (MCI) deve-se à doença de Alzheimer. Este trabalho define um protocolo o qual determina a probabilidade da ocorrência da patologia e classifica pacientes de acordo com essa probabilidade.

2.1.2 Biomarcadores adquiridos por Imagens

Devidos aos avanços na aquisição, qualidade e processamento de imagens neurais dos últimos anos, estudos que fazem uso de biomarcadores baseados em imageamento têm demonstrado melhorias na precisão da detecção da AD (HANYU et al., 2010) (GRAY et al., 2012) (CHEN et al., 2013). O Imageamento por Ressonância Magnética, MRI, e a Tomografia por Emissão de Pósitrons são as técnicas de imageamento mais utilizadas para determinar a progressão da doença. Enquanto imagens PET são mais adequadas a identificar os estágios iniciais da doença, MRI são amplamente utilizadas na detecção da degeneração mais acentuada a qual se faz presente nos estágios mais avançados do AD. A seguir, são descritos os principais tipo de imagens cerebrais utilizadas no diagnóstico da AD e os biomarcadores associados a elas.

Devido à sua natureza não-invasiva e sua capacidade de fornecer imagens de boa resolução espacial e contraste, o imageamento por ressonância magnética (MRI) têm sido amplamente utilizada para esse fim. Assim, diversos estudos têm usado biomarcadores baseados em imagens de Ressonância Magnética Estrutural (*Structured Magnetic Resonance Imaging*, *s-MRI*) para

classificação da AD devido à sua capacidade de descrever atrofia cerebral e alterações no tamanho de tecidos nervosos (LAMA et al., 2017).

De forma similar, o Imageamento por Ressonância Magnética funcional (*Functional Magnetic Resonance Imaging, f-MRI*) pode ser utilizado para detectar a respostas hemodinâmicas relevantes à atividade neural e à conectividade funcional e estrutural (LAMA et al., 2017). Assim, o f-MRI pode ser usado para descrever desordens em todo cérebro a nível de conectividade fazendo dele um biomarcador relevante, junto com o s-MRI para detecção da etapas onde a degeneração cerebral está mais avançada.

2.2 APRENDIZADO DE MÁQUINA

O termo Aprendizado da Máquina ou *Machine Learning* (ML) é amplamente definido como o conjunto de métodos computacionais que usam experiência prévia ou dados de exemplo para, de forma indutiva, produzir previsões ou descrições mais precisas (DIETTERICH, 2009). Uma definição amplamente citada, porém mais formal, estabelece os algoritmos estudados na área do ML como programas de computador cuja performance, medida por \mathbf{P} , é melhorada através da experiência \mathbf{E} com respeito a algum conjunto de tarefas \mathbf{T} (MITCHELL, 1997).

Em termos práticos, o uso de ML baseia-se na construção de um modelo definido sobre alguns parâmetros, onde o aprendizado é realizado por meio da execução de um algoritmo que otimiza estes parâmetros usando dados de treinamento ou experiência passada. Este modelo pode ser ou preditivo, para fazer previsões no futuro, ou descritivo para obter-se conhecimento sobre os dados, ou ambos (ALPAYDIN, 2010).

Assim, o principal objetivo do ML é adaptar modelos (de previsão ou descrição) aos dados. Isto permite que o ML se torne uma ferramenta valiosa ao processo científico pois este, para realizar previsões ou descrições, frequentemente constrói teorias a partir do conhecimento que é extraído de dados que são coletados da observação de experimentos. Segundo Alpaydin (2016), este processo de coleção de dados e construção de teorias/modelos continua até que sejam produzidos modelos e teorias suficientemente explicáveis. Porém, como este tipo de análise de dados é difícil de ser realizada manualmente, uma vez que existem poucas pessoas capazes de fazê-lo e, adicionalmente, o volume disponível de dados atualmente é enorme, torna-se desejável o desenvolvimento de modelos capazes de automatizar tal tarefa e esta é uma preocupação da disciplina de aprendizado de máquina.

O aprendizado de máquina é utilizado em variadas tarefas computacionais onde criar e

programar algoritmos explícitos é impraticável. Exemplos de aplicações incluem filtragem de *spam* (DRUCKER; WU; VAPNIK, 1999), processamento de linguagem natural, reconhecimento óptico de caracteres (*Optical Characters Recognition, OCR*), motores de busca, bioinformática, diagnósticos médicos (LIMA; SILVA-FILHO; SANTOS, 2016b), reconhecimento de fala e de escrita, visão computacional, jogos digitais (GALWAY; CHARLES; BLACK, 2008)(LUCAS, 2008) e locomoção de robôs. Para treinar um sistema inteligente de tal forma que seja capaz de resolver os problemas relacionados com domínio de cada uma dessas aplicações são utilizados diferentes tipos de aprendizado conforme a seguir.

2.2.1 Tipos de Aprendizado

Para que sistemas se tornem inteligentes, é necessário que estes sejam capazes de aprender. Na área de inteligência artificial existem modelos que aprendem por dedução, ou seja, partindo-se de afirmativas lógicas que são combinadas segundo regras estabelecidas para produzir novas afirmativas, e também modelos que aprendem de maneira indutiva, ou seja, através de um conjunto particular de exemplos. O aprendizado indutivo pode ser dividido em três principais tipos de aprendizado caracterizados pelo tipo de sinal ou resposta (*feedback*) obtido ao longo do seu treinamento. São eles: aprendizado supervisionado, aprendizado não-supervisionado e aprendizado por reforço os quais são detalhados a seguir.

2.2.1.1 Aprendizado Supervisionado

No aprendizado supervisionado, o sistema ou agente inteligente tem acesso a conhecimento *a priori* (prévio) estruturado como exemplos que relacionam o estado do ambiente às ações associadas a este estado. Dessa forma, estes exemplos correspondem a pares de entradas X (estado do ambiente) e saídas Y (ações) e a tarefa é aprender o mapeamento da entrada para a saída. Por exemplo, em estatística, regressão é a tarefa de estimar um valor de saída (contínuo) dado um valor específico de entrada (também contínuo) onde é dado um conjunto de pares entrada/saída conhecidos. Em ML, a regressão é um problema de aprendizado supervisionado.

O aprendizado supervisionado também é aplicado em problemas de classificação os quais podem ser vistos como um caso especial de regressão onde a variável de saída Y é um código (valor existente em um conjunto finito de valores) que representa uma classe. Se o número classes é igual a 2, o problema de classificação é denominado binário. Caso o número de classes

seja maior que 2, então o problema de classificação é multiclases.

Cada exemplo ou caso é representado por um vetor de características X_i com $i \in (1, 2, \dots, N)$ (também chamado de vetor de atributos). Cada atributo ou variável x_j com $j \in (1, 2, \dots, M)$, representa um determinado aspecto ou característica (*feature*) do exemplo. De forma geral, existem dois tipos de atributos: numéricos e nominais (também chamado de categóricos). Os atributos numéricos podem ser contínuos, quando representados por números reais ou discretos quando representados por números inteiros. Os atributos nominais correspondem a elementos pertencentes a um conjunto finito de valores não-ordenáveis.

Assim, a definição formal da tarefa de um aprendizado supervisionado é:

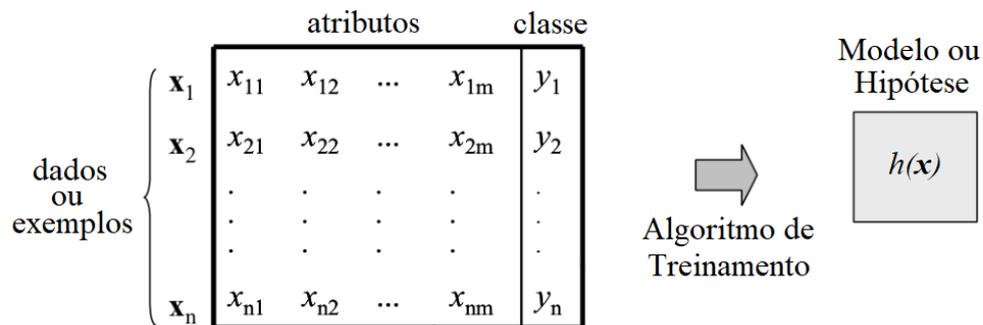
Dado um conjunto de treinamento de N pares de exemplo entrada-saída

$$\Psi = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\},$$

onde cada y_i foi gerado por uma função desconhecida $y = f(x)$, descobrir uma função h que aproxime a verdadeira função f .

A função h é uma hipótese e o aprendizado é uma busca através do espaço de possíveis hipóteses por uma que consiga realizar de forma satisfatória esse mapeamento entre entrada e saída, mesmo com novos exemplos que estão fora do conjunto de treinamento. A Figura 2 apresenta o processo de treinamento aplicado para obtenção do modelo ou hipótese o qual é construído a partir de processo indutivo de aprendizagem supervisionada.

Figura 2 – Modelo Discriminante construído a partir de aprendizagem supervisionada.



Fonte: Adaptado de Alpaydin (2010)

Em outros casos, a função f é estocástica, ou seja, não é estritamente uma função de x , então a hipótese que deve ser aprendida é uma distribuição de probabilidade condicional, $P(Y|x)$. Para medir a precisão de uma hipótese, é utilizado um conjunto de teste formado

por exemplos distintos daqueles existentes no conjunto de treinamento. Uma hipótese é considerada de adequada capacidade de **generalização** quando ela é capaz de prever o valor de y para novos exemplos. No caso onde o algoritmo de treinamento induz uma hipótese que apresente uma baixa taxa de acerto mesmo no subconjunto de treinamento, diz-se que ocorre uma condição de sub-ajustamento ou *underfitting* dos dados com o modelo encontrado. Essa situação pode ocorrer, por exemplo, quando a quantidade de dados de treinamento é insuficiente ou os exemplos de treino disponíveis são pouco representativos ou ainda quando o modelo obtido é muito simples. Por outro lado, quando ocorre do modelo se especializar nos dados utilizados em seu treinamento mas apresentando uma baixa taxa de acerto quando confrontado com os dados de teste (novos dados), tem-se a ocorrência de um superajustamento ou *overfitting*. Assim, encontrar uma hipótese complexa que se adapte de forma exata **a todos** os exemplos de treinamento ou encontrar uma hipótese mais simples que possua boa capacidade de generalização, de forma geral, são objetivos conflitantes (RUSSELL; NORVIG, 2009).

Por fim, na área de ML existe uma preocupação com alguns tipos de problemas que podem estar presente nos dados como dados incorretos (com o rótulo da classe errado), incompletos (faltando valores de alguns atributos) ou contendo ruídos (atributos com erros em seus valores). É importante que os algoritmos de treinamento sejam capazes de minimizar o impacto causado pela utilização de dados com tais problemas.

2.2.1.2 *Aprendizado Não Supervisionado*

No aprendizado supervisionado, o objetivo é aprender um mapeamento entre a entrada e uma saída cujos valores corretos são conhecidos. Porém, no aprendizado **não** supervisionado não existem informações prévias sobre a saída e o agente dispõe apenas dos dados de entrada. Dessa forma, o objetivo é encontrar padrões de regularidade na entrada. Ou seja, o aprendizado não-supervisionado é aplicado quando deseja-se aprender mais sobre os dados.

A tarefa de aprendizado não-supervisionado mais comum é o agrupamento ou *clustering* no qual o objetivo é descobrir agrupamentos potencialmente úteis de exemplos de entrada.

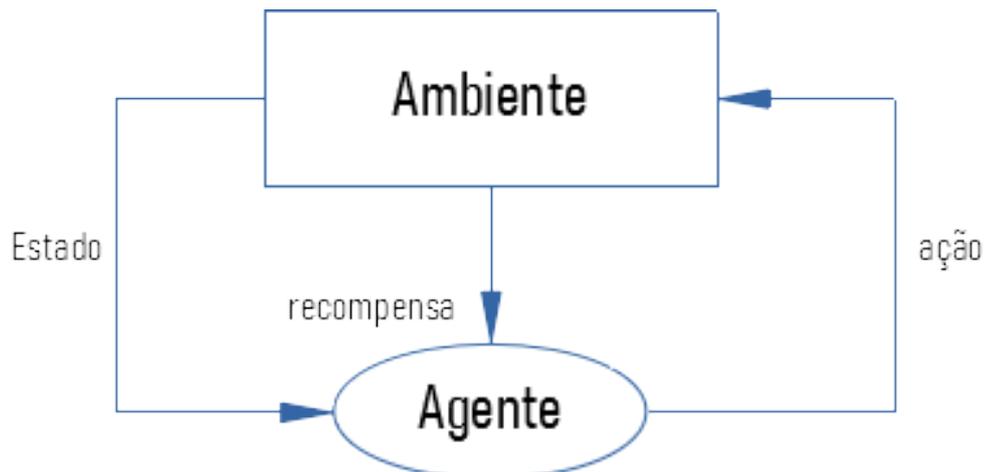
2.2.1.3 *Aprendizado por Reforço*

Em aprendizado por reforço, o agente inteligente é um tomador de decisões sobre quais ações ele deve tomar, em um dado ambiente, para resolver um determinado problema de forma

que uma recompensa ou penalidade é recebida de acordo com o resultado de suas escolhas. Após diversas tentativas, o agente deve aprender a melhor política a qual é sequência de ações que maximizam a recompensa total.

Em qualquer momento, o ambiente está em um certo estado pertencente a um conjunto dos possíveis estados. Dado o estado atual, o agente possui um conjunto de ações possíveis. Uma vez que uma ação é escolhida e tomada, o estado é modificado. Neste processo, é comum o recebimento de pequenas recompensas (ou penalidades) à medida que as ações intermediárias são tomadas. Este processo é ilustrado na Figura 3. Segundo Manning, Raghavan e Schütze (2008), alcançar o objetivo final requer a execução de uma série de ações e uma recompensa rara será recebida geralmente apenas quando a sequência completa é realizada. Por fim, a melhor sequência aprendida corresponde à sequência de ações que possui a maior recompensa acumulativa.

Figura 3 – Interação entre agente e ambiente.



Fonte: Adaptado de Alpaydin (2010)

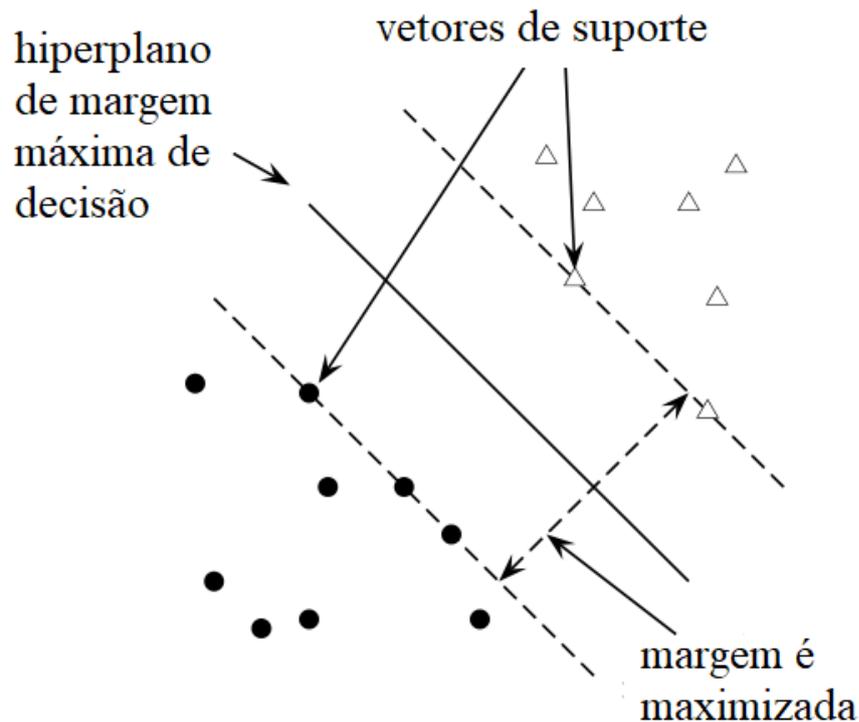
2.3 MÁQUINAS DE VETOR DE SUPORTE

Máquina de Vetor de Suporte é um método de ML para classificação binária baseado em espaço vetorial no qual o objetivo é encontrar uma fronteira de decisão entre as duas classes de tal maneira que seja máxima a sua distância para qualquer ponto no conjunto de treinamento. Esta distância entre a superfície de decisão e os pontos de dados mais próximos é chamada de **margem** do classificador (MANNING; RAGHAVAN; SCHÜTZE, 2008)(p.319).

O método de máquina de vetores de suporte foi inicialmente desenvolvido por Boser, Guyon e Vapnik (1992) em 1992. Este método posteriormente ficou conhecido como SVM de Margem

Rígida *Hard Margin SVM* pois era capaz de trabalhar apenas com dados completamente linearmente separáveis e sem erros (*outliers* ou ruídos). Quando ocorriam erros, ou a margem era reduzida, ou o método falhava em construir uma hipótese de capacidade de generalização adequada. Um modelo gerado por SVM de margem rígida é exibido na Figura 4.

Figura 4 – Um Modelo construído por SVM de Margem Rígida



Fonte: Manning, Raghavan e Schütze (2008)

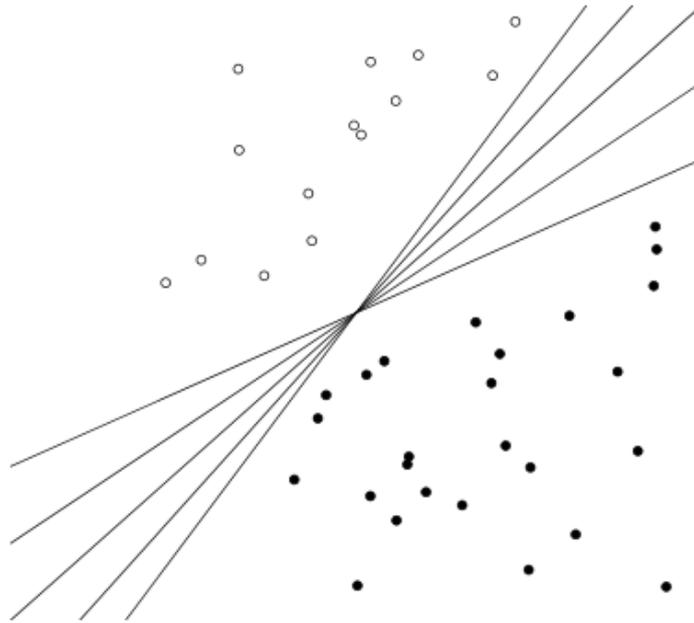
Devido às limitações do SVM de Margem Rígida, foi proposta uma nova versão em (CORTES; VAPNIK, 1995) chamada SVM de Margem Flexível. A seguir, serão apresentadas em mais detalhes a versão original do método como também sua versão aprimorada.

2.3.1 SVMs de Margem Rígida

Para classificação binária, nos casos onde os conjunto de treinamentos são separáveis linearmente, podem ser construídas muitos separadores lineares como é apresentado na Figura 5. Porém, intuitivamente é possível afirmar que uma fronteira de decisão tomada exatamente no meio do espaço compreendido entre os dados das duas classes parece uma melhor opção do que outras que sejam muito próximas dos exemplos de uma ou de ambas as classes. Esta é a ideia central do SVM: buscar a fronteira de separação cuja distância para qualquer um dos pontos de dados que seja maximizada. Isto necessariamente significa que a fronteira de

decisão é especificada totalmente por um subconjunto, usualmente pequeno, de pontos de dados que definem a posição dessa função discriminante. Estes pontos são conhecidos como **vetores de suporte**. A Figura 4 apresenta a margem de separação e os vetores de suporte. Os outros pontos de dados não fazem parte da determinação da superfície de decisão.

Figura 5 – Infinitos separadores lineares separam duas classes linearmente separáveis



Fonte: Manning, Raghavan e Schütze (2008)

A definição formal do SVM pode ser feita por um termo de interceptação b , e um vetor normal ao hiperplano de decisão, \vec{w} , o qual seja perpendicular ao hiperplano. Este vetor \vec{w} é conhecido como vetor peso. Seja o conjunto de treinamento T que contem n instâncias de dados $\mathbf{x}_i \in X$ e seus respectivos rótulos $y_i \in Y$, onde X constitui o espaço de dados e $Y = -1, +1$, o hiperplano discriminante é definido na equação 2.1.

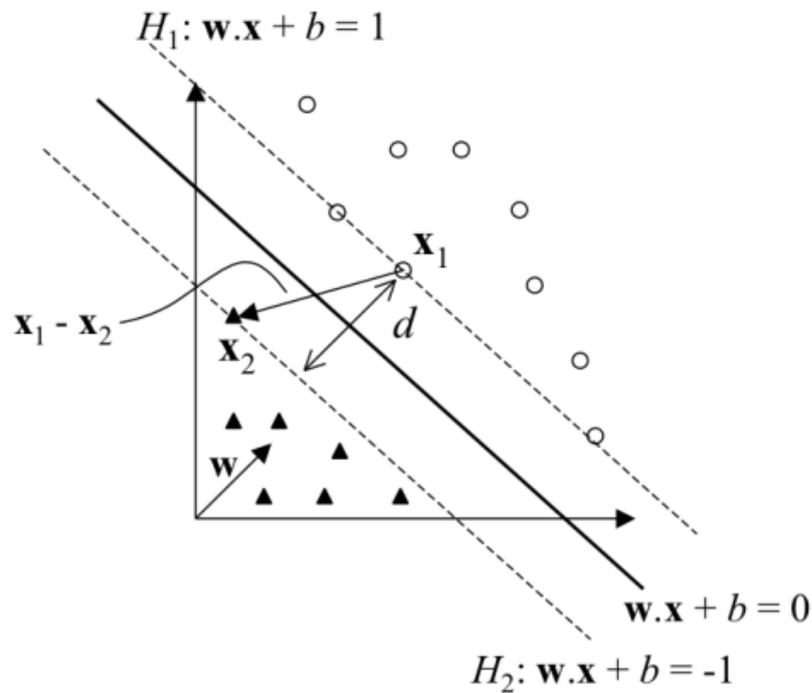
$$f(x) = \mathbf{w} \cdot \mathbf{x} + b = 0 \quad (2.1)$$

O hiperplano definido acima que separa o espaço vetorial dos dados em duas regiões: $\vec{w} \cdot \vec{x} + b \geq 0$ e $\vec{w} \cdot \vec{x} + b \leq 0$ de modo que uma função sinal pode ser utilizada para a realização das classificações ilustrado na Equação 2.2.

$$g(x) = \text{sgn}(f(x)) = \begin{cases} +1 & \text{se } \mathbf{w} \cdot \mathbf{x} + b \geq 0, \\ -1 & \text{se } \mathbf{w} \cdot \mathbf{x} + b \leq 0. \end{cases} \quad (2.2)$$

Sejam \mathbf{x}_1 um ponto no hiperplano $H_1 : \mathbf{w} \cdot \mathbf{x}_1 + b = +1$ e \mathbf{x}_2 um ponto no hiperplano $H_2 : \mathbf{w} \cdot \mathbf{x}_2 + b = -1$, como ilustrado na Figura 6.

Figura 6 – Cálculo da distância entre hiperplanos



Fonte: Lorena e Carvalho (2007)

A projeção de $(x_1 - x_2)$ sobre o vetor de pesos w corresponde à distância d entre os hiperplanos H_1 e H_2 e pode ser calculada a partir da diferença entre $H_1(x_1) - H_2(x_2)$ conforme a Equação 2.4:

$$w \cdot x_1 + b = +1 \quad (2.3)$$

$$- w \cdot x_2 + b = -1$$

$$\frac{w \cdot (x_1 - x_2) = 2}{\quad} \quad (2.4)$$

Multiplicando ambos os lados da igualdade da Equação 2.4 por $\frac{1}{\|w\|}$, é obtido o valor da distância entre os hiperplanos (lado esquerdo da Equação 2.5) a qual corresponde ao dobro do valor da margem.

$$\frac{w \cdot (x_1 - x_2)}{\|w\|} = \frac{2}{\|w\|}. \quad (2.5)$$

O valor obtido, $\frac{2}{\|w\|}$, quando tomado seu máximo, é chamado de margem geométrica, ρ , do classificador, o que é equivalente a afirmar que ρ corresponde à largura máxima da faixa que pode ser desenhada separando os vetores de suporte das duas classes (MANNING; RAGHAVAN; SCHÜTZ, 2008)(p.323). Assim, para maximizar a margem geométrica ρ é preciso encontrar

\mathbf{w} e b tal que:

- $\rho/||\mathbf{w}||$ é máximo e
- Para todo $(\mathbf{x}_i, y_i) \in T, y_i(\mathbf{w}_T \mathbf{x}_i + b) \geq 1$

Como a distância de um dado \mathbf{x}_i até o discriminante é

$$\frac{|\mathbf{w} \cdot \mathbf{x}_i + b|}{||\mathbf{w}||} \quad (2.6)$$

a qual, quando $y_i \in \{-1, +1\}$, pode ser escrita como

$$\frac{y_i(\mathbf{w} \cdot \mathbf{x}_i + b)}{||\mathbf{w}||} \quad (2.7)$$

e esta distância deve ser no mínimo o valor de ρ

$$\frac{y_i(\mathbf{w} \cdot \mathbf{x}_i + b)}{||\mathbf{w}||} \geq \rho, \forall i \quad (2.8)$$

e, fixando ρ como

$$\rho = \frac{1}{||\mathbf{w}||}, \quad (2.9)$$

permanece a restrição $\forall (\mathbf{x}_i, y_i) \in T, y_i(\mathbf{w}_T \mathbf{x}_i + b) \geq 1$ mas a expressão a ser maximizada torna-se

$$\frac{2}{||\mathbf{w}||^2} \quad (2.10)$$

a qual é equivalente a minimizar $\frac{1}{2}||\mathbf{w}||^2$, o que define o problema de otimização como:

$$\arg \min_{\mathbf{w}} \frac{1}{2}||\mathbf{w}||^2 \quad (2.11)$$

Isto corresponde à formulação padrão do SVM como um problema de minimização: Encontrar \mathbf{w} e b tal que:

- $\frac{1}{2}||\mathbf{w}||^2$ é minimizada e
- para todo $(\mathbf{x}_i, y_i) \in T, y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1$

O problema de otimização obtido é quadrático sujeito a restrições lineares. Como a função objetivo é convexa e os pontos que satisfazem a restrição formam um conjunto convexo, este problema possui um único mínimo local (HAYKIN, 2008)(p.273). Problemas desse tipo

podem ser solucionados com a introdução de uma função Lagrangiana cuja solução envolve a construção de um problema *dual* onde um multiplicador de Lagrange α_i é associado a cada restrição $y_i(\mathbf{w}_T \mathbf{x}_i + b) \geq 1$ no problema *primal* para que elas sejam incluídas na função objetivo (MANNING; RAGHAVAN; SCHÜTZE, 2008). Dessa forma, o problema *dual* corresponde a encontrar $\alpha_1, \dots, \alpha_N$ tal que:

$$\begin{aligned} L(\alpha, \mathbf{w}, b) &= \frac{1}{2} \|\mathbf{w}\|^2 - \sum_{i=1}^n \alpha_i (y_i (\mathbf{w} \cdot \mathbf{x}_i + b) - 1) \\ &= \frac{1}{2} \|\mathbf{w}\|^2 - \sum_{i=1}^n \alpha_i (y_i (\mathbf{w} \cdot \mathbf{x}_i + b)) + \sum_{i=1}^n \alpha_i \end{aligned} \quad (2.12)$$

A função Lagrangiana deve ser minimizada, o que implica em minimizar \mathbf{w} e b e maximizar as variáveis α_i . Tem-se, então, um ponto de sela no qual os gradientes de L (Equação 2.12) com respeito a \mathbf{w} e b são zero e também que $\alpha_i \geq 0$ conforme abaixo:

$$\frac{\partial L}{\partial \mathbf{w}} = 0 \Rightarrow \mathbf{w} = \sum_i \alpha_i y_i \mathbf{x}_i \quad (2.13)$$

$$\frac{\partial L}{\partial b} = 0 \Rightarrow \sum_i \alpha_i y_i = 0 \quad (2.14)$$

Inserindo-se as Equações 2.13 e 2.14 na Equação 2.12 a solução é:

$$\arg \max_{\alpha} \sum_j -\frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j (\mathbf{x}_i \cdot \mathbf{x}_j) \quad (2.15)$$

$$\text{Com as restrições: } \begin{cases} \sum_i \alpha_i y_i = 0 \\ \alpha_i \geq 0, \quad \forall i = 1, \dots, n \end{cases} \quad (2.16)$$

Uma vez que o vetor α_i seja encontrado, é possível determinar \mathbf{w} através da equação 2.13. Porém, existem três importantes propriedades da Equação 2.15 que tornam vantajoso utilizar a representação dual. Primeira, a expressão é convexa. Em segundo, os dados entram na expressão apenas na forma de produtos internos de pares de exemplos. Por fim, com o vetor α determinado, percebe-se que a maioria dos α_i são zero, mas aqueles cujo valor é diferente de zero indicam que o correspondente \mathbf{x}_i é um vetor de suporte.

De acordo com Manning, Raghavan e Schütze (2008), a segunda propriedade é também verdade para a equação do discriminante que, uma vez que o α é encontrado, corresponde a:

$$h(\mathbf{x}) = \text{sign} \left[\sum_i \alpha_i y_i (\mathbf{x}_i \cdot \mathbf{x}) + b \right] \quad (2.17)$$

Os vetores de suporte são os exemplos \mathbf{x}_i que satisfazem $y_i(\mathbf{w} \cdot \mathbf{x}_i + b) = 1$ e estão localizados sobre a margem. Este fato pode ser utilizado para calcular b a partir qualquer vetor

suporte usando

$$b = y_i - \mathbf{w} \cdot \mathbf{x}_i. \quad (2.18)$$

Por estabilidade numérica, isto deve ser feito para todos os vetores suportes, tomando um valor médio de b . A função discriminante resultante é, assim, chamada de Máquina de Vetor de Suporte (LORENA; CARVALHO, 2007).

2.3.2 SVMs de Margem Flexível

No mundo real é comum encontrar conjuntos de dados que não sejam linearmente separáveis devido a presença de ruído e também de dados discrepantes ou *outliers*. Assim, nestes casos não é possível obter um hiperplano capaz de separar as classes de forma que não haja erros de classificação. A estratégia, então, é tentar encontrar pelo hiperplano que produzirá o menor erro possível. Para isso, as SVMs de Margem Rígida são estendidas para um novo método que seja capaz de lidar com conjuntos de treinamento mais gerais onde existam alguns dados que possam violar a restrição $y_i(\mathbf{w}_T \mathbf{x}_i + b) \geq 1$. Este método se chama SVM de Margem Flexível ou Suave e é obtido através da introdução de variáveis de folga (*slack variables*), $\xi_i \geq 0$, que são responsáveis por armazenar, para cada dado de exemplo \mathbf{x}_i , um valor correspondente ao desvio da margem cometido no respectivo exemplo de dado. Assim, um valor não-zero para ξ_i permite que o dado \mathbf{x}_i não atenda a restrição de margem a um custo proporcional ao valor de ξ_i .

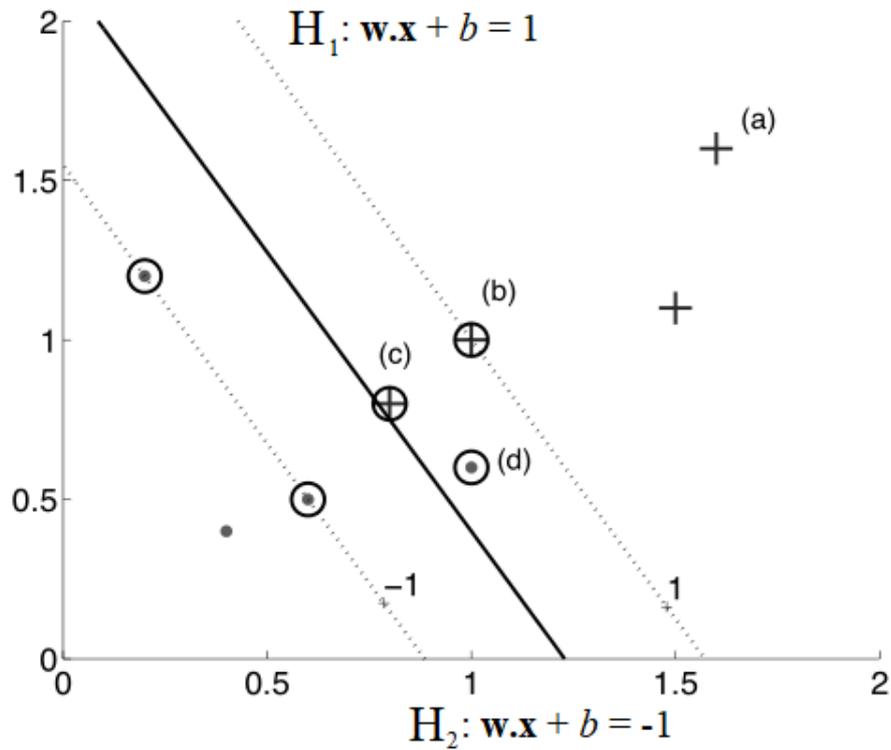
A formulação do SVM de Margem Suave produz uma versão relaxada das restrições impostas ao problema de otimização *primal* conforme a seguir:

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 - \xi_i \quad (2.19)$$

$$\xi_i \geq 0, \quad \forall i = 1, \dots, n \quad (2.20)$$

Esta suavização das margens do classificador linear permite que alguns pontos de exemplo permaneçam entre os hiperplanos H_1 e H_2 que definem as margens (ver Figura 7). Assim, se $\xi_i = 0$, então não há problemas com o dado de exemplo \mathbf{x}_i . Se $0 < \xi_i < 1$, então \mathbf{x}_i é corretamente classificado, embora esteja localizado dentro da margem. Porém, se $\xi_i \geq 1$, então \mathbf{x}_i é classificado erroneamente. O valor $\sum_i \xi_i$ é definido como *soft error* (erro de suavização) o qual é adicionado à Equação 2.11 como um termo de penalidade, definindo um novo problema

Figura 7 – Classificação com restrições relaxadas



Fonte: Alpaydin (2010)

de otimização:

$$\arg \min_{\mathbf{w}} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_i \xi_i \quad (2.21)$$

a qual está sujeita às penalidades definidas nas Equações 2.19 e 2.20.

O parâmetro C é um termo de regularização que provê uma forma de controlar o *overfitting*. Com ele, a escolha do hiperplano ótimo é ajustada para produzir margens mais largas e melhor posicionadas permitindo, assim, uma boa capacidade de generalização do modelo, contudo sem que também ocorram muitos erros na classificação de prováveis *outliers* durante o treinamento.

Adicionando as restrições, o Lagrangiano da Equação 2.12 se torna:

$$\begin{aligned}
L &= \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_i \xi_i - \sum_i \alpha_i [y_i(\mathbf{w} \cdot \mathbf{x}_i + b) - 1 + \xi_i] - \sum_i \mu_i \xi_i \\
&= \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_i \xi_i - \sum_i [\alpha_i y_i (\mathbf{w} \cdot \mathbf{x}_i) + \alpha_i y_i b - \alpha_i + \alpha_i \xi_i] - \sum_i \mu_i \xi_i \\
&= \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_i \xi_i - \sum_i \alpha_i y_i (\mathbf{w} \cdot \mathbf{x}_i) - \sum_i \alpha_i y_i b + \sum_i \alpha_i - \sum_i \alpha_i \xi_i - \sum_i \mu_i \xi_i \\
&= \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_i \xi_i - \sum_i \alpha_i y_i (\mathbf{w} \cdot \mathbf{x}_i) - b \sum_i \alpha_i y_i + \sum_i \alpha_i - \sum_i \alpha_i \xi_i - \sum_i \mu_i \xi_i \\
&= \frac{1}{2} (\mathbf{w} \cdot \mathbf{w}) + (C - \alpha_i - \mu_i) \sum_i \xi_i - \sum_i \alpha_i y_i (\mathbf{w} \cdot \mathbf{x}_i) - b \sum_i \alpha_i y_i + \sum_i \alpha_i \quad (2.22)
\end{aligned}$$

onde μ_i são os novos parâmetros Lagrangianos incluídos para garantir a positividade de ξ_i .

Tomando-se as derivadas parciais com respeito aos parâmetros e igualando-as a zero, tem-se:

$$\frac{\partial L}{\partial \mathbf{w}} = 0 \Rightarrow \mathbf{w} = \sum_i \alpha_i y_i \mathbf{x}_i \quad (2.23)$$

$$\frac{\partial L}{\partial b} = 0 \Rightarrow \sum_i \alpha_i y_i = 0 \quad (2.24)$$

$$\frac{\partial L}{\partial \xi_i} = 0 \Rightarrow C - \alpha_i - \mu_i \quad (2.25)$$

e, por fim, inserindo as Equações 2.23, 2.24 e 2.25 na Equação 2.22, o problema de otimização obtido é:

$$\begin{aligned}
\arg \max_{\alpha_i} L &= \frac{1}{2} (\mathbf{w} \cdot \sum_i \alpha_i y_i \mathbf{x}_j) - \sum_i \alpha_i y_i (\mathbf{w} \cdot \mathbf{x}_i) + \sum_i \alpha_i \\
&= \frac{1}{2} (\sum_i \alpha_i y_i (\mathbf{w} \cdot \mathbf{x}_j) - \sum_i \alpha_i y_i (\mathbf{w} \cdot \mathbf{x}_i)) + \sum_i \alpha_i \\
&= \sum_i \alpha_i - \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y_i y_j (\mathbf{x}_i \cdot \mathbf{x}_j) \quad (2.26)
\end{aligned}$$

que é a mesma formulação apresentada para as SVMs de Margem Rígida, mas com a restrição dos α_i limitados pelo valor de C :

$$\begin{cases} \sum_i \alpha_i y_i = 0 \\ 0 \leq \alpha_i \leq C, \quad \forall i = 1, \dots, n. \end{cases} \quad (2.27)$$

Seja α^* a solução do problema *dual* e \mathbf{w}^* , b^* e ξ_i^* as soluções do problema *primal*, o vetor \mathbf{w}^* , continua sendo calculado a partir da Equação 2.13. As variáveis ξ_i podem ser calculadas pela Equação 2.28:

$$\xi_i^* = \max \left\{ 0; 1 - y_i \sum_{j=1}^n y_j \alpha_j^* \mathbf{x}_j \cdot \mathbf{x}_i + b^* \right\} \quad (2.28)$$

A variável b^* é novamente calculada utilizando-se α^* , ou seja, levando em conta apenas os valores de $y_i(\mathbf{w} \cdot \mathbf{x}_i + b^*) - 1 + \xi_i$ dos vetores de suporte (quando $\alpha_i = 0$) e tomando-se o valor médio de b^* conforme a Equação 2.29:

$$\alpha_i^* [(y_i \mathbf{w} \cdot \mathbf{x}_i + b^*) - 1 + \xi_i] = 0 \quad (2.29)$$

obedecendo à restrição

$$(C - \alpha^*) \xi_i^* = 0. \quad (2.30)$$

2.4 REDES NEURAS CONVOLUCIONAIS

Redes convolucionais, também conhecidas como Redes Neurais Convolucionais, ou CNNs (Convolutional Neural Networks) são um tipo especializado de rede neural para processar dados que têm uma topologia tipo grade ou matricial como dados de uma dimensão tomados ao longo do tempo ou imagens. As CNNs possuem esse nome por utilizarem a operação matemática chamada convolução, que é um tipo especializado de operação linear, no lugar de uma multiplicação de matrizes ao menos em alguma camada (GOODFELLOW; BENGIO; COURVILLE, 2016).

2.4.1 A Operação de Convolução

A operação convolução pode ser entendida como um tipo de valor médio de uma função de entrada que é ponderado através de uma segunda função, conhecida como função *kernel*. Dessa forma, a operação de convolução pode ser definida formalmente conforme a seguir:

Seja x uma função de entrada (ou *input*) tomada sobre valores contínuos de um argumento t e seja w uma função peso cujo valor é tomado sobre a diferença entre t e um segundo argumento, a , então a convolução s (ou *output*) da função de entrada x ponderada pela função *kernel* w é definida como:

$$s(t) = (x * w)(t) = \int_{-\infty}^{+\infty} x(a)w(t - a)da. \quad (2.31)$$

Caso x e w sejam definidas apenas sobre valores inteiros do argumento t , então a operação é chamada convolução discreta e é definida como:

$$s(t) = (x * w)(t) = \sum_{a=-\infty}^{\infty} x(a)w(t - a). \quad (2.32)$$

Em aprendizado de máquina, a entrada é usualmente um vetor de dados multidimensional e o *kernel* é geralmente um vetor multidimensional de parâmetros que são adaptados pelo algoritmo de treinamento, onde estes dois vetores multidimensionais são chamados de tensores.

Nos casos onde a entrada corresponde à uma imagem 2D, a convolução ocorre sobre dois eixos e normalmente o *kernel* será também bidimensional. Dessa forma, a convolução S aplicada a uma imagem I utilizando um *kernel* K é definida como:

$$S(i, j) = (I * K)(i, j) = \sum_m \sum_n I(m, n)K(i - m, j - n). \quad (2.33)$$

onde os valores máximos de m e n correspondem à largura e a altura da imagem I .

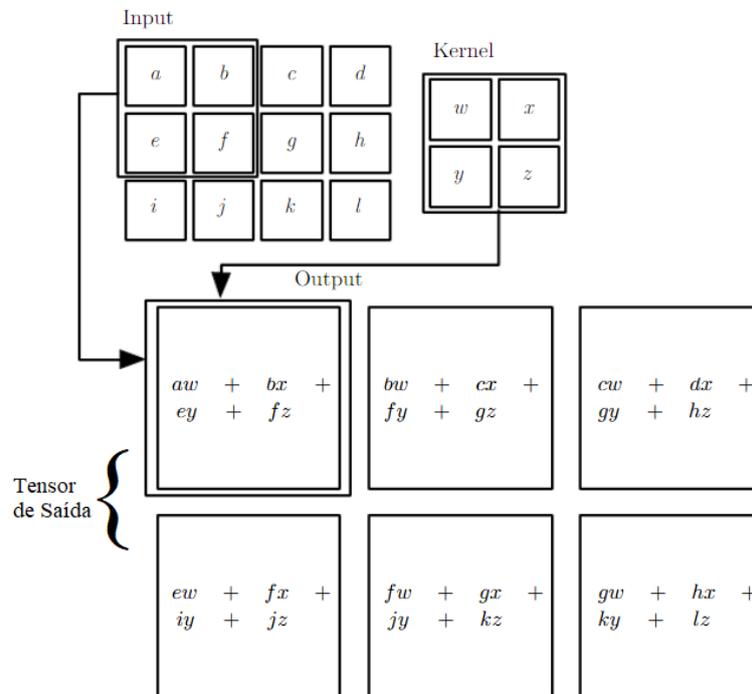
Porém, é comum fazer uso da propriedade de comutatividade da operação de convolução para obter-se uma convolução equivalente, porém com uma menor variação no intervalo de valores válidos m e n . Assim, é equivalente a Equação 2.33 ser reescrita produzindo a Equação 2.34:

$$S(i, j) = (I * J)(i, j) = (K * I)(i, j) = \sum_m \sum_n I(i - m, j - n)K(m, n). \quad (2.34)$$

Contudo, na Equação 2.34, apesar da comutatividade ser útil para escrever provas, ela não é usualmente uma propriedade importante para uma implementação de redes neurais artificiais. Assim, muitas bibliotecas de redes neurais artificiais implementam um função relacionada chamada de correlação-cruzada a qual é a mesma convolução mas sem inverter o Kernel. Dessa forma, em ML é comum a utilização de uma função alternativa (relacionada com a operação de convolução) chamada **correlação cruzada** onde não ocorre a inversão do *kernel*, (GOODFELLOW; BENGIO; COURVILLE, 2016). A correlação cruzada é definida pela Equação 2.35:

$$S(i, j) = (I * J)(i, j) = (K * I)(i, j) = \sum_m \sum_n I(i + m, j + n)K(m, n). \quad (2.35)$$

Apesar da mudança, na área de ML é comum implementar a correlação cruzada e chamá-la de convolução. Dessa forma, ao adotar-se essa nomenclatura, é importante especificar-se sobre a inversão ou não do *kernel* quando for relevante. Na Figura 8, é apresentado um exemplo de convolução (sem inversão do *kernel*) aplicada a um tensor de duas dimensões. Neste exemplo, a saída foi restrita a posições onde o *kernel* repousa inteiramente dentro da imagem. As setas e caixas indicam como o elemento superior mais à esquerda do tensor de saída é construído aplicando-se o kernel à região superior-esquerda correspondente ao tensor de entrada

Figura 8 – Exemplo de convolução sem inversão de *kernel*

Fonte: Adaptado de Goodfellow, Bengio e Courville (2016)

2.4.2 Características das redes neurais convolucionais

Modelos de aprendizado de máquina baseados em CNNs, quando comparados àqueles baseados em Redes Neurais tradicionais, se beneficiam de alguns importantes aspectos resultantes da utilização da operação de convolução, conforme a seguir:

Interações esparsas

Em redes neurais tradicionais, em uma dada camada, cada unidade de saída interage com cada uma das unidades de entrada. Isso ocorre devido à multiplicação de matriz efetuada entre o vetor de entrada e o vetor de parâmetros onde cada parâmetro descreve a interação entre uma dada unidade de entrada e uma dada unidade de saída. Porém, a utilização de um *kernel* menor que a entrada permite que as CNNs possuam **interações esparsas**, uma vez que a convolução produz uma unidade de saída utilizando apenas um pequeno subconjunto de elementos da entrada. No exemplo da Figura 8, um *kernel* 2x2 é aplicado a uma entrada bidimensional 3x4 produzindo uma saída 2x3 onde cada unidade da saída interage com apenas quatro unidades da entrada. Este subconjunto de unidades de entrada que influenciam o valor

de um elemento de saída específico é chamado de **Campo Receptivo** deste elemento de saída.

Compartilhamento de parâmetros

O compartilhamento de parâmetros se refere ao uso de um mesmo parâmetro por mais de uma função em um modelo. Em redes neurais tradicionais, cada elemento da matriz de parâmetros é utilizado apenas um vez no cálculo da saída a partir da multiplicação entre a matriz dos dados de entrada e a matriz de parâmetros. Em redes convolucionais, cada elemento da matriz de parâmetros, *kernel*, é utilizado no cálculo de cada um dos elementos de saída (ver exemplo apresentado na Figura 8).

Representações equivariantes

Uma função f é dita equivariante a outra função g quando $f(g(x)) = g(f(x))$ ou seja, equivariante significa que se uma entrada sofre uma dada modificação, a saída é modificada da mesma maneira. Assim, a operação de convolução é considerada equivariante a uma operação de translação, pois a saída produzida pela função *kernel* para um dado subconjunto de unidades da entrada é a mesma, independente da posição desse subconjunto dentro da entrada. Esta propriedade pode ser útil se o interesse sobre uma dada característica é definido sobre sua presença na entrada, e não a respeito de sua localização. Contudo, a operação de convolução não é naturalmente invariante a outras transformações tais como mudanças na escala ou rotação de uma imagem. Outros mecanismos são necessários para manipular estes tipos de transformações (GOODFELLOW; BENGIO; COURVILLE, 2016).

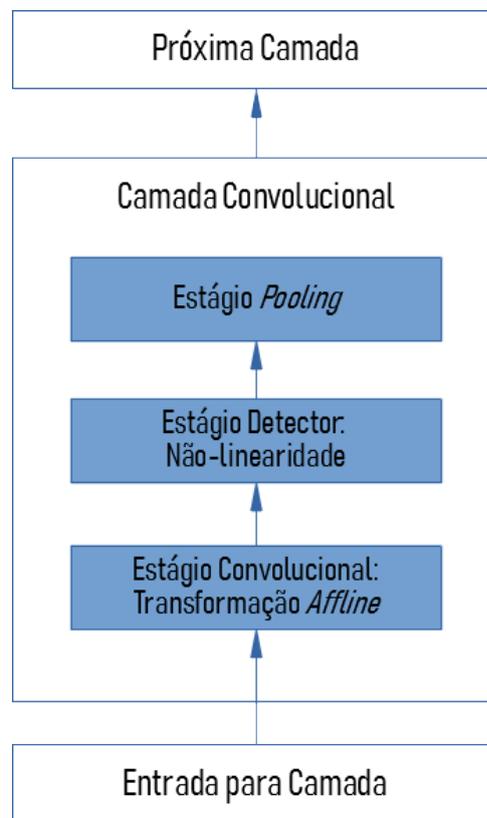
2.4.3 Arquitetura de uma rede convolucional

Uma CNN é normalmente composta de diversas camadas convolucionais que extraem características dos dados recebidos na entrada e as entregam a uma camada totalmente conectada (FC, *Full Connected*) a qual é responsável por realizar a classificação ou predição.

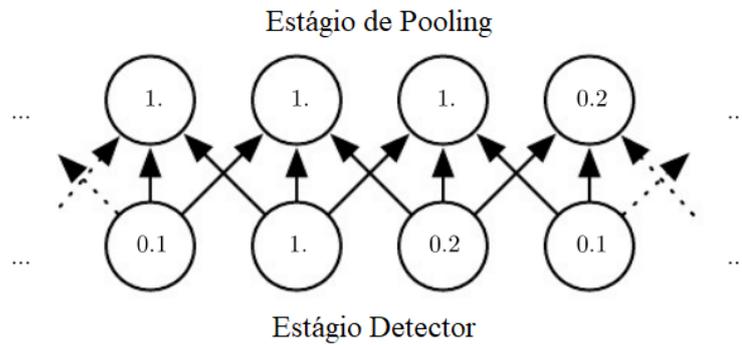
A estrutura típica de uma camada convolucional, apresentada na Figura 9, consiste tipicamente de três estágios. No primeiro estágio, são realizadas diversas convoluções em paralelo para produzir um conjunto de ativações lineares. No segundo estágio, cada ativação linear é

processada por uma função de ativação não-linear. Este estágio é também chamado de estágio detector. No terceiro estágio, uma função de combinação, ou *pooling*, é utilizada para modificar a saída da camada. A função de *pooling* substitui o valor da saída de uma rede em uma certa localização, por uma estatística resumida das saídas vizinhas. A Figura 10 apresenta um exemplo de função de *pooling* chamada *maxpooling* a qual reporta a saída máxima dentro de uma vizinhança retangular. Por fim, a camada FC é tipicamente definida como uma rede neural tradicional a qual é treinada usando os atributos extraídos pelas camadas convolucionais.

Figura 9 – Estágios que compõem uma camada típica de uma rede neural convolucional



Fonte: Adaptado de Goodfellow, Bengio e Courville (2016)

Figura 10 – Função *Maxpooling*

Fonte: Adaptado de Goodfellow, Bengio e Courville (2016)

2.5 CONSIDERAÇÕES

Este capítulo destacou os principais conceitos e teorias necessários à compreensão da proposta realizada neste trabalho. Foram apresentados conceitos relacionados com os métodos computacionais utilizados na proposta e também as questões que envolvem a Doença de Alzheimer e sua classificação a partir de soluções apoiadas computacionalmente.

No próximo capítulo (3), é apresentada uma breve revisão narrativa de trabalhos de classificação de pacientes com Alzheimer baseado em neuroimageamento. São comentados os aspectos metodológicos envolvendo cada estudo e também os resultados obtidos por cada proposta. Além disso, também são realizadas observações sobre a capacidade de generalização e reprodutibilidade dos estudos.

3 TRABALHOS RELACIONADOS

Neste capítulo são apresentados os trabalhos que foram considerados mais relacionados a esta pesquisa de doutorado, dentre as publicações mais recentes.

3.1 ESTADO DA ARTE

Ben-Ahmed et al. (2014) propõem um esquema onde Fusão Tardia (*late fusion*) é utilizada para combinar duas modalidades diferentes de dados provenientes de volumes MRI: assinatura visual da atrofia do hipocampo e o volume do CSF o qual preenche o espaço causado por tal atrofia. No processo de extração de atributos, as imagens são inicialmente normalizadas em dimensão e orientação segundo o modelo MNI. Para a extração do volume relativo ao hipocampo, é utilizado o atlas AAL para marcar quais pontos (*voxels* do volume total serão usados para reconstruir uma nova imagem volumétrica da região de interesse. No final do processo, o método *Circular Harmonic Functions, CHF*) é aplicado nestas imagens volumétricas para extrair atributos que são submetidos a uma classificação binária por meio de um classificador SVM com kernel RBF (*Radial Basis Function*). A extração dos atributos de CSF é baseada na contagem de *voxels* de baixa intensidade (escuros) existentes na região do hipocampo que são identificados a partir de um limiar de intensidade pré-definido. A classificação dos dados provenientes do volume CSF é realizada por um classificador Bayesiano. As saídas probabilísticas de ambos classificadores são tomadas e, juntas, formam vetores de atributos de segunda ordem, com dimensão igual a 2, que são utilizados para treinar um terceiro classificador, também SVM. Os experimentos foram realizados utilizando imagens MRI estruturais da base ADNI de 1.5T e ponderadas em T1 de 218 indivíduos sendo, dos quais, 35 pacientes AD, 72 controles (NC) e 111 indivíduos com MCI. Também foram realizados experimentos utilizando-se imagens MRI ponderadas em 3T de 21 indivíduos controles e 16 pacientes diagnosticados com AD provenientes de uma base de dados francesa chamada *Bordeaux dataset*. Os autores não mencionam se foi utilizado algum método de balanceamento das classes minoritárias (nos experimentos com ambas as bases). O estudo também não cita a aplicação de algum método para contornar possíveis problemas relacionados a utilização de um baixo número de elementos e uma grande quantidade de atributos no treinamento do classificador Bayesiano assim como no primeiro SVM utilizado. Sobre a base de dados francesa, os autores não informam se foi

realizada alguma subamostragem ou se a base era, de fato, composta de apenas 37 indivíduos. O estudo obteve 87% (ADNI) e 85% (*Bourdeaux*) de acurácia na classificação binária AD vs NC, e 78,22% (ADNI) e 72,23% para MCI vs NC e MCI vs AD, respectivamente. O estudo observou um aumento médio de 9% na acurácia após utilizar a fusão tardia.

Uma arquitetura baseada em redes profundas CNN, *Convolution Neural Network*, é utilizada por Farooq et al. (2017) para realizar a classificação multi-classe dos estágios do Alzheimer. Neste estudo, são extraídos 355 volumes MRI da base ADNI, no formato Nifti, provenientes de 33 pacientes com AD, 22 pacientes diagnosticados como LMCI, 49 com MCI e 45 indivíduos controle saudáveis (NC). Estes volumes são transformados em fatias 2D para compor o conjunto de imagens que será usado nos experimentos. São propostos três modelos onde ambos utilizam rede convolucionais. A acurácia obtida pelos três modelos foi: 98,88% para a abordagem com GoogLeNet, 98,01% com ResNet de 18 camadas e 98,14% para ResNet com 152 camadas. Contudo, o trabalho tem difícil reprodutibilidade pois não há informações sobre a quantidade de fatias utilizadas nem quais fatias exatamente foram selecionadas para os experimentos. O estudo também possui um problema metodológico com o balanceamento das classes onde novos elementos são sinteticamente criados a partir de cópias produzidas pelo espelhamento horizontal das imagens originais. Esta metodologia pode não ser adequada visto que, dependendo do *kernel* utilizado, a operação de convolução torna-se invariável a giro (*flip*) horizontal (GONZALEZ; WOODS, 2008). Além disso, dado que o cérebro humano possui um relevante nível de simetria entre os hemisférios, dependendo da orientação das imagens originais, a produção de novas imagens por *flip* horizontal das existentes pode gerar novas imagens de alto nível de semelhança. Portanto, o *flip* horizontal pode produzir dentro do conjunto de teste diversas cópias de imagens utilizadas no treinamento. Assim, visto que não há informações específicas sobre os *kernels* utilizados, os modelos propostos no estudo podem não ter validade alguma pois os mesmos dados são utilizados no treinamento também são utilizados durante o teste.

Gorji e Haddadnia (2015) utilizaram duas redes para realizar a classificação binária entre as fases do Alzheimer: uma rede LVQ, *Learning Vector Quantization*, e uma rede PRNN, *Pattern Recognition Neural Network*. Ambas são treinadas por 3 diferentes métodos: *Levenberg-Marquardt (LM) back-propagation*, *Resiliente Back Propagation (RB)*, e *Scaled Conjugate Gradient (SCG)*. Os melhores resultados foram obtidos pela rede PRNN treinada pelo algoritmo SCG. Na classificação binária entre AD e NC, o modelo proposto obteve 97,27%, entre as classes AD e MCI obteve 94,88% e entre MCI e NC conseguiu 95,55% de acurácia.

A extração de atributos é realizada manualmente por meio de Pseudo Momentos de Zernick (BHATIA; WOLF, 1954). Foram utilizados dados de imagens MRI da base ADNI1 de um conjunto de 500 indivíduos composto de 180 pacientes AD, 172 MCI e 188 indivíduos saudáveis (NC). O estudo dividiu, através de sorteio aleatório, os dados de exemplo da seguinte forma: 50% para treinamento, 25% para validação e 25% para teste. Não foi utilizada validação cruzada. Durante o treinamento, a técnica *Early Stop* foi utilizada para evitar *overfit*. Os autores, contudo, não descrevem de forma clara a metodologia utilizada para a aquisição, a partir dos volumes Nifti, das imagens 2D necessárias ao método de extração de atributos utilizado. Dessa forma, estes problemas dificultam a replicação desse estudo.

O objetivo do estudo realizado por Salvatore, Cerasa e Castiglioni (2018) foi treinar um classificador para distinguir pacientes que precisam de tratamento (AD + MCI Progressivo) daqueles que não precisam (NC + MCI Estável). Para isso, foi utilizada uma abordagem multimodalidade onde foram utilizados dados provenientes de imagens como também dados neuropsicológicos. Os experimentos foram realizados utilizando dados da base ADNI de 200 indivíduos, dos quais 50 foram diagnosticados com AD, outros 50 como MCI progressivo, 50 MCI estáveis e 50 Controles normais (NC). As imagens utilizadas já foram adquiridas com a correção *Grandwarp* e a correção de não-uniformidade B1. Adicionalmente, os autores optaram por realizar nas imagens outros procedimentos como recorte, re-orientação e remoção do crânio para, depois, segmentá-las de acordo com os tecidos de massa cinzenta (GM) e de massa branca (WM) e suavizá-las em seguida. O conjunto de atributos foi construído a partir de atributos extraídos e ranqueados por PCA (*Principal Component Analysis*) e FDR, e atributos extraídos através de *Partial Least Squares*, *PLS*. A seleção e otimização dos atributos foi realizada por um método *wrapper* (empacotamento de alguns atributos) validado por uma abordagem de validação cruzada aninhada (*Nestled Cross Validation*). A classificação foi realizada através de um classificador SVM com um *kernel* linear implementado pelos autores. Por fim, fazendo uso apenas dos dados de imagem, a acurácia máxima obtida foi de 72% na classificação binária entre o grupo que não precisa de tratamento (formado por MCI estável e NC) e o outro grupo de pacientes que precisam de tratamento (formado por MCI progressivo e AD). Nos experimentos que fizeram uso de dados de imagem e de dados neuropsicológicos, o melhor resultado obtido foi de 85% de acurácia. Mesmo utilizando dados observados em diferentes momentos (6, 18 e 24 meses) antes da confirmação do diagnóstico, o modelo proposto obteve resultados semelhantes. Contudo, o estudo não apresenta detalhes sobre a subamostragem realizada na base ADNI que permitiu selecionar exatamente 50 indivíduos de cada classe de

forma que não ficou claro se as propriedades estatísticas foram preservadas nesse processo ou se ocorreu algum viés de seleção.

Asim et al. (2018) propuseram uma abordagem multimodalidades na qual são utilizados dados de imagens MRI e FDG-PET (fluoro-deoxi-glicose PET) para realizar a classificação binária entre AD e MCI, entre MCI e NC e entre AD e NC. São utilizados dados de 300 indivíduos distribuídos entre os três grupos (AD, MCI e NC), com cada um contendo exatamente 100 indivíduos que foram tomados aleatoriamente da base ADNI (o estudo não menciona se as propriedades estatísticas foram mantidas após este processo de subamostragem). Foram realizados pré-processamento de suavização e normalização espacial (MRI E FDG-PET) e segmentação (apenas MRI) além de correções de intensidade (N3). As imagens MRI foram utilizadas para aquisição de informações estruturais para avaliar a densidade de massa cinzenta, de massa branca e de líquido cefalorraquidiano (CSF), enquanto as imagens PET foram utilizadas para avaliar o metabolismo cerebral. A extração de atributos das imagens é baseada em atlas, ou seja, leva em conta as região anatômicas em análise seguindo um mapa pré-determinado. O método PCA (*Principal Component Analysis*) é aplicado para reduzir a dimensionalidade do vetor de atributos construído, selecionando os atributos mais promissores. O classificador utilizado de melhor desempenho foi o SVM com *kernel* de função de base radial (RBF), cujos parâmetros foram otimizados através de *grid search*. O treinamento foi validado através de validação cruzada LOO (*leave-one-out*). O modelo proposto obteve 94% de acurácia na classificação entre AD e NC, 76,5% entre MCI e NC e 75,5% entre MCI e AD.

Promteangtrong et al. (2015) apresentam uma revisão sistemática das potenciais aplicações clínicas das técnicas tradicionais e novas de ressonância magnética (RM) na avaliação de pacientes com doença de Alzheimer (DA), incluindo ressonância magnética estrutural, ressonância magnética funcional, imagem por tensão de difusão e imagem por transferência de magnetização. De acordo com Tondelli et al. (2012), a análise de sMRI é capaz de detectar DA até 10 anos antes do diagnóstico clínico. As duas características patológicas mais prevalentes associadas à demência são atrofia cortical, incluindo atrofia do lobo temporal medial e alterações vasculares. A ressonância magnética estrutural (sMRI) é importante para o diagnóstico diferencial da DA devido à sua capacidade de visualizar padrões específicos de atrofia no cérebro. A atrofia do hipocampo, um biomarcador comum de ressonância magnética, tem sido considerada um critério-chave para o diagnóstico de DA.

Tosun et al. (2013) tentaram identificar uma assinatura de neuroimagem preditiva de amiloidose cerebral como uma ferramenta de triagem para identificar indivíduos com comprome-

timento cognitivo leve (MCI) que são mais propensos a ter altos níveis de amiloidose cerebral ou a não amiloide. De acordo com Tosun et al. (2013), as variações anatômicas da forma preditiva de amiloidose cerebral em MCI abraçaram um padrão espacial característico conhecido por alta vulnerabilidade à patologia da doença de Alzheimer, incluindo o lobo temporal medial, córtices de associação temporal-parietal, cíngulo posterior, precuneus, hipocampo, terminais da amígdala, caudado e fórnice/estria. O desempenho na predição de positividade para beta-amilóide ($A\beta+$) usando sMRI e dados genéticos (ApoE) foi significativamente superior ao usar apenas um dos biomarcadores.

Promteangtrong et al. (2015) também destacam que a atrofia geral do lobo temporal médio está comumente associada a outros tipos de demências e não apenas à DA. Levando em conta essa questão, outras pesquisas começaram a tentar identificar regiões específicas do lobo temporal médio cujas alterações estruturais são mais adequadas para a detecção de DA (HYMAN et al., 1984; GÓMEZ-ISLA et al., 1996; BOBINSKI et al., 1999; ARENDT et al., 2015; MIKKONEN et al., 1999; GIANNAKOPOULOS et al., 2003; ŠIMIĆ et al., 2000; LI et al., 2021). Vários estudos apontam para descobertas patológicas, principalmente perda neuronal, relacionadas à DA (HYMAN et al., 1984; GÓMEZ-ISLA et al., 1996; BOBINSKI et al., 1999; ARENDT et al., 2015; MIKKONEN et al., 1999; GIANNAKOPOULOS et al., 2003; ŠIMIĆ et al., 2000; LI et al., 2021). Essas perdas parecem estar localizadas mais predominantemente no córtex entorrinal (HYMAN et al., 1984; GÓMEZ-ISLA et al., 1996; BOBINSKI et al., 1999; ARENDT et al., 2015; MIKKONEN et al., 1999; GIANNAKOPOULOS et al., 2003; ŠIMIĆ et al., 2000; LI et al., 2021). Khan et al. (2014) usaram uma variante de ressonância magnética funcional de alta resolução (fMRI) que pode mapear defeitos metabólicos em pacientes e modelos de camundongos para abordar questões básicas sobre a fisiopatologia do córtex entorrinal. O córtex entorrinal é dividido em regiões funcionalmente distintas, o córtex entorrinal medial (MEC) e o córtex entorrinal lateral (LEC). Os autores exploraram os recursos de alta resolução da variante de fMRI para perguntar se algum deles foi afetado em pacientes com doença de Alzheimer pré-clínica. O córtex entorrinal tem sido implicado nos estágios iniciais da doença de Alzheimer. É caracterizada por alterações na proteína tau e nos fragmentos clivados da proteína precursora de amilóide (APP). Khan et al. (2014) usou uma variante de ressonância magnética funcional de alta resolução (fMRI) que pode mapear defeitos metabólicos em pacientes e modelos de camundongos para abordar questões básicas sobre a fisiopatologia do córtex entorrinal. Empiricamente, o córtex entorrinal parece ser bem visualizado a partir de imagens do plano sagital (KHAN et al., 2014; HYMAN et al., 1984; GÓMEZ-ISLA et al., 1996; BOBINSKI et al., 1999; ARENDT et al., 2015; MIKKONEN et al.,

1999; GIANNAKOPOULOS et al., 2003; ŠIMIĆ et al., 2000; LI et al., 2021).

Fang et al. (2017) extraíram dados de 25 regiões corticais em cada hemisfério, pois o objetivo era avaliar se realizar a seleção de recursos e o processo de treinamento para os hemisférios esquerdo e direito do cérebro separadamente traria melhorias significativas.

Uma rede hierárquica usando regiões promissoras já conhecidas de todo o cérebro e um esquema de classificação foram construídos por Liu et al. (2016) para identificar regiões mais adequadas. Ramírez et al. (2018) aplicaram o método baseado em um teste ANOVA para remover características não informativas para classificação.

Li, Cheng e Liu (2017) propõem um método de classificação baseado na combinação de redes convolucionais 3D multimodelo para aprender os vários recursos de imagens cerebrais de RM. Os autores construíram uma rede neural convolucional 3D profunda para transformar a imagem de RM em recursos de alto nível mais compactos. Eles também usaram autoencoders convolucionais 3D multiescala para extrair outros recursos de imagem. Esses recursos foram combinados com as camadas superiores totalmente conectadas para classificação de imagens no diagnóstico de DA. Nesse processo, os autores evitam a segmentação cerebral. Para avaliar sua proposta, os autores usaram imagens cerebrais de RM ponderadas em T1 de 428 indivíduos, incluindo 199 pacientes com DA e 229 controles normais (NC) do banco de dados da Alzheimer's Disease Neuroimaging Initiative (ADNI1). Eles alcançaram uma precisão de 88,31% para classificação AD e uma AUC (área sob a curva ROC) de 92,73%.

Basaia et al. (2019) propuseram um algoritmo de aprendizado profundo que prevê o diagnóstico individual da doença de Alzheimer (DA) e comprometimento cognitivo leve que se converterá em DA (c-MCI) com base em uma única ressonância magnética estrutural do cérebro em corte transversal. Redes neurais convolucionais (CNNs) foram aplicadas em imagens 3D ponderadas em T1 de ADNI e indivíduos recrutados em instituto próprio (407 controles saudáveis [HC], 418 AD, 280 c-MCI, 533 MCI estável [s-MCI]). O desempenho da CNN foi testado na distinção de AD, c-MCI e s-MCI. Os autores conseguiram alcançar altos níveis de precisão em todas as classificações, com as maiores taxas alcançadas nos testes de classificação AD vs HC usando apenas o conjunto de dados ADNI (99%) e o conjunto de dados ADNI + não-ADNI combinado (98%). CNNs discriminaram c-MCI de pacientes s-MCI com uma precisão de até 75% e nenhuma diferença entre imagens ADNI e não-ADNI. Os autores afirmam que sua proposta tem potencial para acelerar a adoção da RM estrutural na prática clínica do neurodiagnóstico.

Long et al. (2017) propuseram um método de aprendizado de máquina para discriminar

pacientes com DA ou comprometimento cognitivo leve (MCI) de idosos saudáveis e prever a conversão da DA em pacientes com MCI por meio da computação e análise das diferenças morfológicas regionais do cérebro entre os grupos. Eles mediram a distância entre cada par de sujeitos a partir de um registro difeomorfo simétrico, seguido por um algoritmo de incorporação e uma abordagem de aprendizado para classificação. O método proposto obteve acurácia de 96,5% na diferenciação de DA leve de idosos saudáveis tendo como região de interesse (ROI) toda a substância cinzenta cerebral ou lobo temporal, 91,74% na diferenciação de MCI progressivo de idosos saudáveis e 88,99% na classificação MCI progressivo versus MCI estável com amígdala ou hipocampo como ROI.

Farooq et al. (2017) propuseram um classificador de 4 vias para classificar Alzheimer (DA), comprometimento cognitivo leve (MCI), comprometimento cognitivo leve tardio (LMCI) e pessoas saudáveis. Eles usaram o conjunto de dados ADNI em um sistema baseado em unidade de processamento gráfico de alto desempenho. Eles realizaram a classificação multiclasse da doença. A técnica proposta resulta em uma precisão de previsão de 98,8%.

Korolev et al. (2017) propuseram modelos de aprendizado profundo baseados em VoxCNN e ResNet para extração de recursos de imagens de ressonância magnética para diagnóstico de DA. Eles usaram o conjunto de dados ADNI de varreduras cerebrais de ressonância magnética estrutural 3D. A precisão média da amostra para o diagnóstico de DA varia de 62-79% para VoxCNN e de 59-80% para ResNet.

Maqsood et al. (2019) propuseram uma técnica baseada no aprendizado de transferência para classificar imagens de MRI ajustando uma rede convolucional AlexNet pré-treinada. A arquitetura é treinada e testada sobre as imagens segmentadas pré-processadas (Matéria Cinzenta, Substância Branca e Líquido Espinhal Cerebral) e não segmentadas para classificação binária e multiclasse. O desempenho foi avaliado sobre o conjunto de dados Open Access Series of Imaging Studies (OASIS). Os autores alcançaram a melhor acurácia geral de 92,85% para classificação multiclasse sobre imagens de RM sem segmentação prévia.

3.2 CONSIDERAÇÕES

Neste Capítulo foi apresentado o estado da arte dos métodos para apoio ao diagnóstico da doença de Alzheimer e de demências por meio da análise de imagens de ressonância magnética. No Capítulo 4 é apresentada a proposta de pesquisa deste trabalho. O Capítulo 5 apresenta um quadro comparativo entre os resultados dos estudos comentados acima e os resultados

obtidos pelo trabalho descrito nesta Tese.

4 PROPOSTA, MATERIAIS E MÉTODOS

Neste capítulo é apresentada a metodologia proposta para a pesquisa, para seleção automática de fatias e planos de visualização de volumes de imagens biomédicas, tendo como estudo de caso o apoio ao diagnóstico da doença de Alzheimer e de déficit cognitivo leve a partir de neuroimagens de ressonância magnética.

4.1 REPRESENTAÇÃO DE IMAGENS

A escolha pela Ressonância Magnética (RM) foi realizada pela maior atenção dada a ela e por suas características não invasivas, assim como a possibilidade de analisar a atrofia cortical e do lobo temporal médio e também o alargamento dos ventrículos.

Um exame de imageamento por ressonância magnética produz um volume ou imagem volumétrica. Este volume é estruturado na forma de um agrupamento tridimensional de diversos *voxels*, os quais equivalem aos *pixels* de imagens bidimensionais. O acesso aos dados de uma imagem volumétrica pode ser realizado através da leitura de todos os *voxels* pertencentes a um plano que corte o volume e que seja paralelo a quaisquer um dos três planos do corpo humano (axial, sagital e coronal). Assim, a imagem volumétrica pode ser também interpretada como um empilhamento, sob um mesmo eixo, de diversas (tipicamente até 255) fatias ou imagens bidimensionais.

Com o objetivo de acelerar os experimentos, uma nova base de dados foi construída a partir de atributos das fatias bidimensionais que compoem cada uma das imagens volumétricas utilizadas. Dessa forma, seja $V = \{v_1, v_2, \dots, v_M\}$, um conjunto de M volumes MRI, uma nova base de dados T é construída de forma que, para cada volume v_m , com $m \in \{1, 2, \dots, M\}$, é produzida uma tabela $t_m \in T$, onde cada j -ésima linha é um vetor de atributos u_j composto de descritores de forma, cor e textura, extraídos da j -ésima fatia de v . Devido ao seu uso amplamente consolidado na literatura, foram utilizados como descritores de forma os 25 Pseudo Momentos de Zernike e, como descritores de textura e cor as 13 primeiras estatísticas de Haralick (calculadas para cada uma das quatro direções 2D). Assim, na construção do conjunto T são extraídos ao todo $(13 \cdot 4) + 25 = 77$ atributos para cada fatia de cada volume. Assim, seja l_m o somatório do total de fatias existentes em cada um dos três planos do corpo humano,

a dimensão de cada $t_m \in T$ é:

$$\dim t_m = l_m.77 \quad (4.1)$$

4.2 MODELAGEM DA OTIMIZAÇÃO

Neste trabalho foi utilizada uma modelagem baseada em computação evolucionária onde a busca pelo melhor conjunto de fatias é modelada como um problema de otimização numérica. Neste tipo de modelagem, tipicamente é gerada uma população inicial composta de soluções candidatas que sofrem pequenas variações incrementais ao longo do tempo com o objetivo de produzir soluções mais adequadas ou aptas à resolução do problema. Dessa forma, este processo evolucionário exigea definição de uma estrutura ou representação para as soluções candidatas de acordo com o problema. Adicionalmente, também é necessário definir uma função objetivo responsável por avaliar numericamente uma dada solução candidata.

4.3 MODELAGEM DOS CANDIDATOS À SOLUÇÃO

Neste trabalho, um candidato a solução corresponde a um conjunto específico de fatias

$$s = (p, i, n), \quad s \in S \quad (4.2)$$

onde

- $n \in [1, N]$ é o número de fatias consecutivas existentes em s ,
- $i \in \{0, 1, \dots, (i_p - n)\}$ é a posição da primeira fatia,
- $p \in \{0, 1, 2\}$ é plano do corpo humano onde estão localizadas as n fatias,
- i_p é o maior valor possível para a posição de uma fatias localizada no p -ésimo plano, considerando todos volumes da base,
- S é o conjunto de todos os candidatos à solução, também conhecido como População do algoritmo evolucionário
- N é o número máximo consecutivos de fatias definidas por um candidato a solução. Por restrições de tempo de processamento, neste trabalho ficou estabelecido $N = 20$.

4.4 FUNÇÃO OBJETIVO

A avaliação do uso de um determinado conjunto de fatias requer a definição da função de transformação

$$g(T, s) = W, \quad (4.3)$$

cujo objetivo é produzir uma base intermediária W , derivada de $T = \{t_1, t_2, \dots, t_M\}$, onde cada $t_m \in T$ produz um w_m que é um vetor construído por concatenação dos atributos de imagem de todas as n fatias definidas em s . Por ser um importante fator de risco para DA, na construção de w_m o valor da idade do paciente também é concatenado com cada vetor $w_m \in W$, onde

$$\dim w_m = 77 \cdot n + 1. \quad (4.4)$$

Assim, o valor da função objetivo f é definido como:

$$f(W, c) = Acc, \quad (4.5)$$

onde Acc é a acurácia geral obtida por um determinado classificador c que foi treinado usando um conjunto de vetores de atributos W para o diagnóstico de DA e MCI. Para maximizar o valor de f , foi aplicado o algoritmo genético para representação com inteiros.

4.5 PIPELINE DE APRENDIZADO DE MÁQUINA

O pipeline de aprendizado de máquina realizado por este trabalho é apresentado a seguir.

4.5.1 Formato de instâncias de treinamento e teste

Na abordagem proposta, avaliar um determinado agrupamento de fatias equivale a fazer uma seleção de dados específicos que serão utilizados para treinar um modelo para o diagnóstico de DA. As instâncias utilizadas por este modelo terão uma dimensionalidade igual ao número de fatias neste agrupamento multiplicado pelo número n de atributos de textura e forma calculados para cada fatia. Portanto, o número de características de um determinado agrupamento de fatias irá variar de 78, onde há apenas uma fatia no agrupamento de fatias, a 1541, nos casos do agrupamento de fatias é composto por 20 fatias. Para evitar a maldição da dimensionalidade, foi aplicado um método de seleção de recursos que é descrito a seguir.

Os experimentos de extração de características consideraram dois cenários: (1) com pré-processamento, extraindo a caixa craniana dos volumes por meio de segmentação não supervisionada (automatizada) por meio de ferramenta BET (Brain Extraction Tool); (2) sem este pré-processamento.

4.5.2 Seleção de características e redução de dimensionalidade

Para acelerar o desempenho do algoritmo de classificação e evitar o *overfitting* devido ao risco de lidar com grande número de características, foi aplicada a técnica de Análise de Componentes Principais, PCA.

A normalização de características (*feature scaling*) é um método usado para padronizar o intervalo de variação das características no aprendizado de máquina. Seu uso é obrigatório, pois a redução de dimensionalidade feita pelo PCA seria enviesada pelas diferenças entre as variâncias dos atributos. Neste trabalho, os dados de imagem foram padronizados para média zero e variância unitária usando $\hat{x} = \frac{x - \bar{x}}{\sigma}$, onde x , \bar{x} e σ são o valor original da característica, sua média e seu desvio padrão, respectivamente.

4.5.3 Balanceamento de classes

Para construir um modelo imparcial, o balanceamento de classes é uma etapa importante do pipeline de aprendizado de máquina. Neste trabalho, empregamos o método SMOTE para construir instâncias sintéticas para classes minoritárias a fim de balancear conjuntos de treinamento e teste (CHAWLA et al., 2002).

4.5.4 Validação do modelo

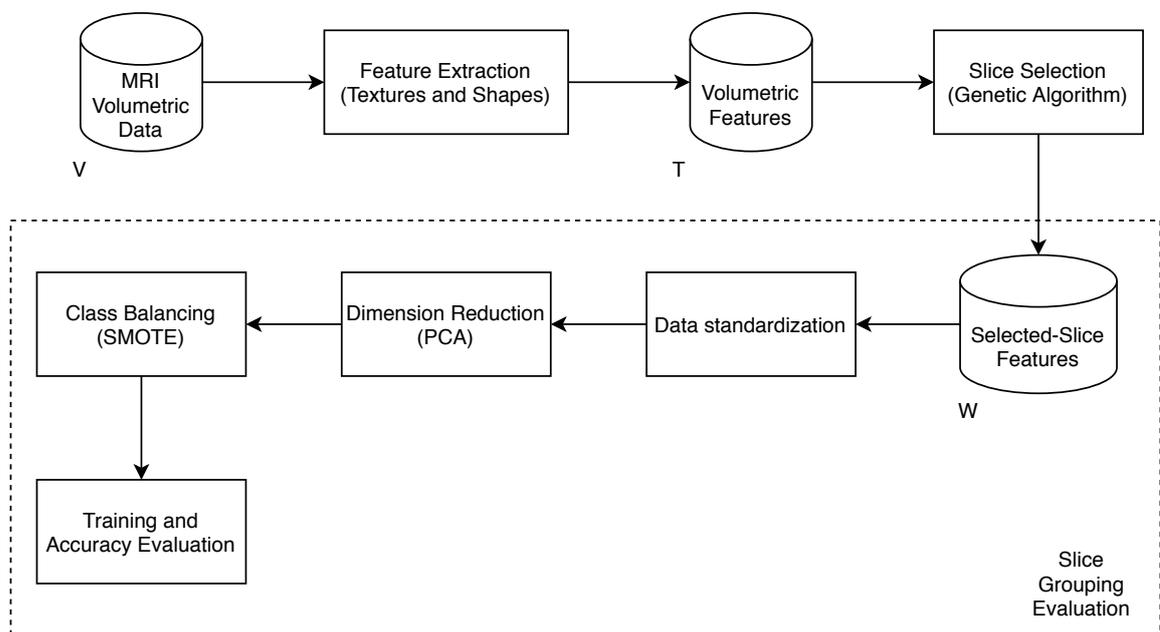
A etapa de validação visa avaliar a capacidade de generalização do modelo. Neste estudo, o *setup* de validação foi feito usando uma abordagem de validação cruzada de 10 subconjuntos.

O pipeline de aprendizado de máquina aplicado para avaliar um determinado agrupamento de fatias é apresentado no fluxograma da Figura 11. Neste processo, todos os volumes de MRI do conjunto de dados são aplicados para o estágio de extração de atributos, no qual cada fatia é representada por atributos de textura e forma. Os cortes são determinados de acordo com a visão definida: axial, coronal ou sagital. Em seguida, as fatias relevantes são

selecionadas usando um algoritmo genético, reduzindo o vetor de características associado a cada volume à concatenação de características de textura e forma para cada fatia selecionada. Após a padronização, as características mais relevantes são selecionadas pela aplicação do método PCA. O classificador é treinado usando uma versão de conjunto de dados balanceada obtida usando aumento de dados com SMOTE. A Figura 12 ilustra como se dá o processo de classificação de acordo com a proposta.

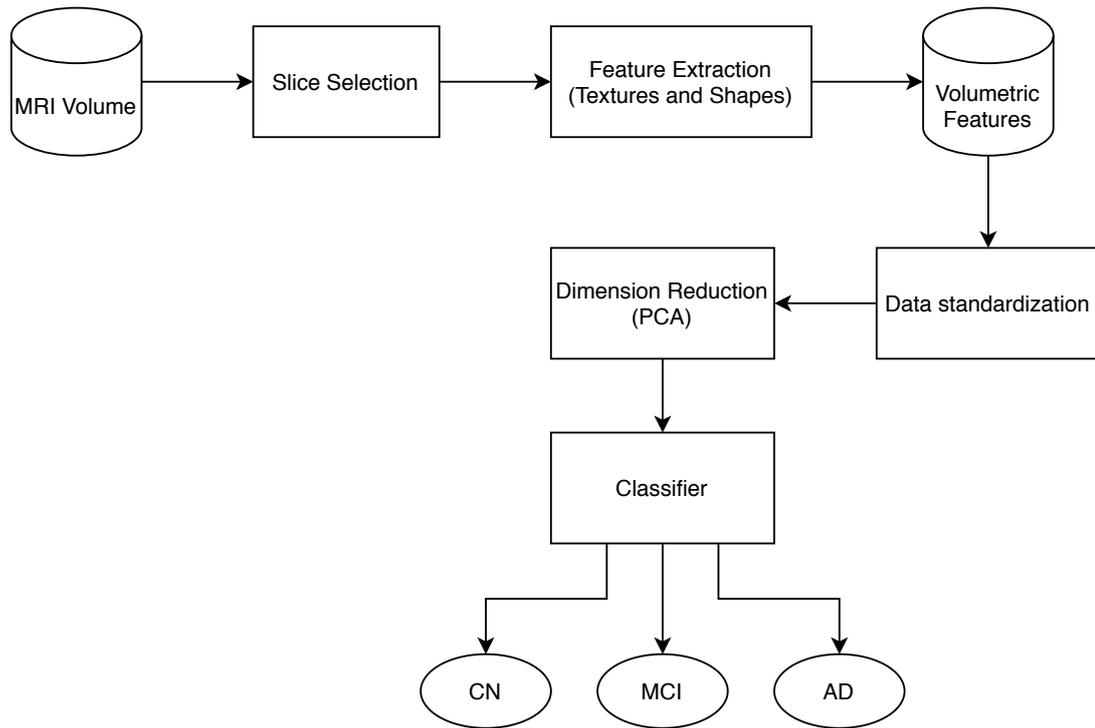
As fatias mais relevantes foram determinadas pelo algoritmo genético na fase de pipeline de aprendizado de máquina. A etapa de redução de dimensão também foi definida na etapa anterior. O classificador dá suporte às informações diagnósticas: se o volume de entrada da RM está associado a um indivíduo saudável (CN), a um indivíduo com comprometimento cognitivo leve ou a um indivíduo com doença de Alzheimer. Os experimentos consideraram dois cenários: (1) com pré-processamento, extraindo a caixa craniana dos volumes por meio de segmentação não supervisionada; (2) sem pré-processamento.

Figura 11 – Proposta: Fluxograma de aprendizado de máquina aplicado para avaliar um agrupamento específico de slices.



Fonte: O Autor

Figura 12 – Uso: classificação de volume de acordo com nossa proposta.



Fonte: O Autor

4.6 PROJETO DE EXPERIMENTOS

Os seguintes algoritmos de classificação foram utilizados para realizar os experimentos: Árvores de Decisão (CART), Análise Discriminante Linear (LDA), Regressão Logística (LR), Naive Bayes (NB), Floresta Aleatória (RF) com 100 árvores e Máquina de Vetor de Suporte (SVM) com kernel linear e $C = 0,1$. O número de vizinhos do classificador kNN e o número de árvores (ou número de estimadores) de Random Forest foram definidos empiricamente. A tabela 1 apresenta os classificadores utilizados, seus respectivos parâmetros ajustados manualmente com seus valores correspondentes.

Em aplicações de otimização baseadas em ferramentas de busca meta-heurística, é importante que a função objetivo, ou seja, a função que orienta o processo de busca e otimização, tenha a menor complexidade computacional possível. Assim, um dos critérios para escolha dos classificadores e suas respectivas configurações foi apresentar complexidade computacional relativamente baixa. Com exceção do Support Vector Machine, todos os métodos têm complexidade de treinamento relativamente baixa. No entanto, dado que SVMs geralmente apresentam bom desempenho de classificação em problemas biomédicos relativamente complexos, o parâmetro C foi definido como 0,1, permitindo margens relativamente suaves para a

Tabela 1 – Parâmetros definidos para os classificadores empregados

#	Classificador	Parâmetros
0	CART	-
1	kNN	$k = 5$
2	LDA	-
3	LR	-
4	NB	-
5	RF	$n = 100$
6	SVM	kernel linear, $C = 0, 1$

Fonte: (SOUZA et al., 2021)

função discriminante. Foi utilizado o kernel linear, ou seja, o kernel mais simples possível, mas ainda capaz de gerar bons resultados, dependendo da natureza dos conjuntos de treinamento e teste.

Adicionalmente, como o comportamento estatístico dos conjuntos de treinamento e teste, ou seja, das fatias dos volumes cerebrais, não é conhecido, a natureza das arquiteturas classificadoras busca refletir as possibilidades. A Análise Discriminante Linear constrói limites de decisão lineares, no caso de problemas linearmente separáveis. A regressão logística fornece um limite de decisão de complexidade logística. O classificador Naive Bayes assume que os atributos das imagens das fatias são variáveis aleatórias independentes e normalmente distribuídas. LDA e NB assumem, portanto, conjuntos de treinamento e teste de relativa simplicidade em relação ao limite de decisão. A Support Vector Machine com kernel linear fornece um limite de decisão de complexidade quadrática, resultante do kernel linear combinado com a decisão por limite linear na saída da máquina de aprendizado, o que resulta em limites de decisão polinomiais de grau dois.

As Árvores de Decisão CART assumem que boas classificações podem ser obtidas se os limites de decisão puderem ser definidos por meio de um sistema de regras. Random Forests consistem em conjuntos de árvores de decisão onde cada árvore é treinada com o conjunto de treinamento composto por vetores de atributos representados com menos atributos que o total. Assim, cada árvore de decisão modela um especialista virtual que vê o todo a partir de seu próprio ponto de vista. A decisão de classificação da Random Forest é tomada por votação. A classe com mais votos é a decisão do classificador. Foram escolhidas 100 árvores para a Random Forest. O classificador Random Forest foi escolhido como classificador para lidar com limites de decisão que são muito mais complexos do que os de uma única árvore de decisão.

Por fim, também optou-se por usar o classificador k vizinho mais próximo, kNN, para modelar conjuntos de treinamento e teste que são difíceis de generalizar. Os classificadores kNN são treinados apenas armazenando amostras de cada classe. A decisão é tomada por votação considerando a classe mais frequente dos k vizinhos mais próximos do vetor de atributos de entrada. Os classificadores kNN são usados quando a falta de conhecimento sobre o comportamento estatístico dos conjuntos de dados é considerada grande. Esses classificadores geralmente modelam sistemas de raciocínio baseados em casos. O uso de cinco vizinhos ($k = 5$) foi definido empiricamente.

Para o processo de otimização, escolhemos um algoritmo genético para selecionar as melhores fatias para classificação de acordo com a acurácia. Os algoritmos genéticos são bastante conhecidos e fornecem parâmetros relativamente intuitivos para serem ajustados. Por esta razão, escolhemos este algoritmo de otimização meta-heurística.

Para cada um dos sete classificadores utilizados, a busca pelo algoritmo genético foi realizada 50 vezes, o que produziu um total de 350 experimentos. Em cada uma dessas execuções, o melhor indivíduo encontrado e seu valor de acurácia correspondente foram registrados na classificação multiclasse dos exames entre Saudável (Cognitivamente Normal, CN), paciente com Comprometimento Cognitivo Leve (CCL) ou Doença de Alzheimer (DA).

Além disso, o banco de imagens sofreu outros dois cortes de acordo com o sexo dos pacientes. Essa divisão do conjunto de imagens a serem utilizadas no treinamento da máquina de aprendizagem é motivada pelas evidências na literatura (MIELKE; VEMURI; ROCCA, 2014; ELBEJJANI et al., 2015; SAMPEDRO et al., 2015; FARRER et al., 1997; ALTMANN et al., 2014; HÖRDER et al., 2018) que mostram que o desenvolvimento e os impactos da Doença de Alzheimer mostram-se diferentes entre os sexos masculino e feminino.

Com isso, todos os experimentos descritos acima foram refeitos utilizando apenas exames de ressonância magnética de pacientes do sexo feminino e, em seguida, utilizando apenas imagens volumétricas de pacientes do sexo masculino. Assim, para cada um dos três cenários produzidos pelo corte de gênero (masculino, feminino e ambos os gêneros), foram realizados um total de 350 experimentos, totalizando 1050 execuções do algoritmo evolutivo.

Por fim, a busca por conjuntos de fatias consecutivas mais promissoras foi realizada para conjuntos com no máximo 20 fatias consecutivas. Essa restrição foi definida para limitar o tempo de processamento necessário para avaliar a adequação de um determinado conjunto de fatias consecutivas.

4.7 BASE DE DADOS

Os dados usados neste trabalho foram obtidos do banco de dados da Alzheimer's Disease Neuroimaging Initiative (ADNI) (<adni.loni.usc.edu>). A ADNI foi lançada em 2003 como uma parceria público-privada, liderada pelo investigador principal Michael W. Weiner, MD. O objetivo principal do ADNI tem sido testar se a ressonância magnética (MRI), a tomografia por emissão de pósitrons (PET), outros marcadores biológicos e a avaliação clínica e neuropsicológica podem ser combinadas para medir a progressão do comprometimento cognitivo leve (CCL) e precoce Doença de Alzheimer (DA).

O projeto ADNI foi lançado em outubro de 2004, envolvendo contribuições de \$67 milhões do Instituto Nacional do Envelhecimento, 13 empresas farmacêuticas e duas fundações (WEINER et al., 2013; WEINER et al., 2010). Este projeto foi desenvolvido para encontrar biomarcadores mais sensíveis e precisos para a detecção precoce e rastreamento da DA. Em seguida, o projeto ADNI-1 analisou milhares de exames cerebrais, perfis genéticos e biomarcadores de sangue e líquido cefalorraquidiano (WEINER et al., 2013; WEINER et al., 2010). A pesquisa incluiu 400 pacientes com diagnóstico de CCL, 200 indivíduos com DA precoce e 200 idosos controles. Os pesquisadores do ADNI-1 empregaram medidas de imagem cerebral, incluindo ressonância magnética estrutural e PET (WEINER et al., 2013; WEINER et al., 2010).

O projeto ADNI foi estendido em 2009 durante a fase ADNI GO (ADNI-2), que avaliou a coorte ADNI1 existente junto com 200 novos participantes com comprometimento cognitivo leve precoce (EMCI) (WEINER et al., 2013; WEINER et al., 2015; WEINER et al., 2017). Esta fase foi projetada para examinar biomarcadores em um estágio inicial da doença (WEINER et al., 2013; WEINER et al., 2015; WEINER et al., 2017).

A fase atual do projeto, ADNI-3, começou em 2016, com um objetivo expandido de determinar as relações entre as características de biomarcadores clínicos, cognitivos, de imagem, genéticos e bioquímicos em todo o espectro da DA (WEINER et al., 2013; WEINER et al., 2015; WEINER et al., 2017). ADNI3 adiciona varreduras cerebrais que detectam emaranhados de proteína tau (tau PET), um indicador chave da doença (WEINER et al., 2013; WEINER et al., 2015; WEINER et al., 2017). O ADNI-3 inclui cientistas em 59 centros de pesquisa nos Estados Unidos e Canadá e adicionará centenas de novos indivíduos MCI, indivíduos com DA leve e controles idosos (WEINER et al., 2013; WEINER et al., 2015; WEINER et al., 2017).

Para avaliações cognitivas, globais, funcionais e comportamentais, o protocolo ADNI3 empregou os seguintes instrumentos (WEINER et al., 2017):

▪ *Avaliações Cognitivas:*

- Alzheimer's Disease Assessment Scale-Cognitive (ADAS-Cog13): é uma escala estruturada administrada pessoalmente por um examinador que avalia a memória (recordação de palavras, reconhecimento de palavras), raciocínio (seguir comandos), linguagem (nomeação, compreensão), orientação, práxis ideacional (colocar carta em envelope) e práxis construtiva (copiar desenhos geométricos). Classificações da linguagem falada, compreensão da linguagem, dificuldade em encontrar palavras e capacidade de lembrar as instruções do teste também são obtidas (ROSEN; MOHS; DAVIS, 1984).
- American National Adult Reading Test (ANART): estima a inteligência verbal pré-mórbida (VIQ) em pacientes com demência, exigindo que os pacientes leiam e pronunciem corretamente 50 palavras irregulares que não seguem as regras comuns de fonografia e ortografia (GROBER; SLIWINSKI; KOREY, 1991; NELSON; O'CONNELL, 1978).
- Testes de Fluência de Categoria: uma medida de fluência verbal em que o participante é solicitado a gerar exemplos das categorias semânticas (animais) em tentativas sucessivas de um minuto (BUTTERS et al., 1987).
- Teste do Desenho do Relógio: testa a função de construção visuoperceptual. Eficaz para discriminar participantes com DA e idosos normais (CAHN et al., 1996).
- Cogstate Brief Battery (CBB): uma breve bateria cognitiva computadorizada (10-15 minutos) desenvolvida pela Cogstate (Cogstate Ltd. New Haven, CT, EUA) que mede atenção, velocidade de processamento de informações, memória de trabalho e aprendizado. A teia testes baseados são acessados usando um site seguro com credenciais de login exclusivas que não exigem nenhuma informação de saúde protegida (MARUFF et al., 2009; MARUFF et al., 2013).
- Teste de Memória Lógica I e II (recordação imediata e tardia de parágrafo): uma medida de memória episódica modificada da Wechsler Memory Scale-Revised (WMS-R). A recordação livre de um conto é eliciada imediatamente após ser lida em voz alta para o participante e novamente após um intervalo de trinta minutos (MANUAL, 1987).

-
- Mini-Mental State Examinations (MMSE): um instrumento de triagem breve e relativamente comum para estudos de medicamentos para a doença de Alzheimer. A escala MMSE avalia orientação, memória, atenção, concentração, nomeação, repetição, compreensão e capacidade de criar uma frase e copiar dois pentágonos sobrepostos (BRAVO; HÉBERT, 1997; FOLSTEIN; ROBINS; HELZER, 1983; BRUCKI et al., 2003; MITCHELL, 2017).
 - Montreal Cognitive Assessment (MoCA): uma breve avaliação cognitiva de 30 pontos projetada para detectar participantes no estágio MCI de disfunção cognitiva (NASREDDINE et al., 2003; FREITAS et al., 2013). Este instrumento demonstrou ter sensibilidade e especificidade adequadas em ambientes clínicos para detectar suspeita de CCL. Seu desempenho é acompanhado para determinar sua capacidade de diferenciar entre os três grupos diagnósticos no ADNI3.
 - Multi-Lingual Naming Test (MINT): um teste de nomeação de imagens de objetos projetado para incluir itens que são comparáveis em inglês, espanhol, mandarim e hebraico (GOLLAN et al., 2012). Ele substitui o Boston Naming Test no Uniform Data Set dos centros AD financiados pela NIA porque o BNT contém itens que não têm o mesmo nível de dificuldade para falantes de espanhol e inglês.
 - Teste de Aprendizagem Verbal Auditiva: uma tarefa de aprendizagem em lista, que avalia vários parâmetros cognitivos associados à aprendizagem e à memória (VAKIL; BLACHSTEIN, 1993; MALLOY-DINIZ et al., 2007).
 - Teste de Trilha A e B: progrida de um quebra-cabeça numérico de conectar os pontos para uma alternância mais desafiadora entre ordem alfa e numérica. Ambas as Partes A e B dependem de habilidades visuomotoras e de varredura perceptiva. A Parte B requer considerável flexibilidade cognitiva na mudança de conjuntos de números para letras sob pressão de tempo também (REITAN, 1958; TOMBAUGH, 2004).
- *Avaliações globais, funcionais e comportamentais:*
- Pontuação Isquêmica de Hachinski Modificada: um breve questionário com perguntas sobre histórico médico, sintomas cognitivos e características do AVC, conforme relatado por um parceiro de estudo e informado pelo exame neurológico e estudos de neuroimagem (ROSEN et al., 1980).

-
- Índice de Mudança Cognitiva (CCI): um instrumento de auto-relato de 20 itens e relatório do parceiro de estudo (informante). Os itens avaliam principalmente a memória com alguma cobertura de alterações executivas e de linguagem. Uma pontuação total de memória CCI é usada para definir o grupo sMCI e as informações do participante e do parceiro de estudo (informante) podem ser combinadas para produzir uma pontuação total para memória ou cognição geral (RATTANABANNAKIT et al., 2016; RISACHER et al., 2015).
 - Clinical Dementia Rating Scale (CDR): descreve cinco graus de comprometimento no desempenho em cada uma das 6 categorias de funcionamento cognitivo, incluindo memória, orientação, julgamento e resolução de problemas, assuntos da comunidade, lar e hobbies e cuidados pessoais (MORRIS, 1993; HUGHES et al., 1982; HUANG et al., 2021; KIM et al., 2017).
 - Questionário de Avaliação de Atividades da Vida Diária / Funcional (FAQ): uma entrevista com um cuidador ou parceiro qualificado, um participante é avaliado em sua capacidade de realizar dez atividades complexas da vida diária (PFEFFER et al., 1982).
 - Financial Capacity Instrument – Short Form (FCI-SF): uma medida de desempenho de 15 minutos das habilidades financeiras diárias que usa variáveis de desempenho e tempo para conclusão (MARSON, 2015). Em pesquisas anteriores, as variáveis de desempenho e tempo de FCI-SF discriminaram entre controles positivos e negativos de amiloide, bem como controle vs MCI vs grupos de DA leve (MARSON, 2015).
 - Medição da Cognição Cotidiana (ECog): avalia comprometimento funcional muito leve que pode ocorrer no CCL. O ECog é um questionário avaliado por informantes, mas os participantes serão solicitados a preencher uma versão autorreferida do mesmo questionário (FARIAS et al., 2008). ECog correlaciona-se bem com medidas estabelecidas de estado funcional e cognição global, mas apenas fracamente com idade e educação.
 - Inventário Neuropsiquiátrico (NPI): um instrumento multi-item, confiável e bem validado para avaliar a psicopatologia na DA com base em uma entrevista com um cuidador ou parceiro de estudo qualificado (definido neste estudo como tendo contato direto 10 horas/semana) (CUMMINGS, 1997; CUMMINGS et al., 1994). Avalia a frequência e gravidade de delírios, alucinações, disforia, ansiedade, agitação ou

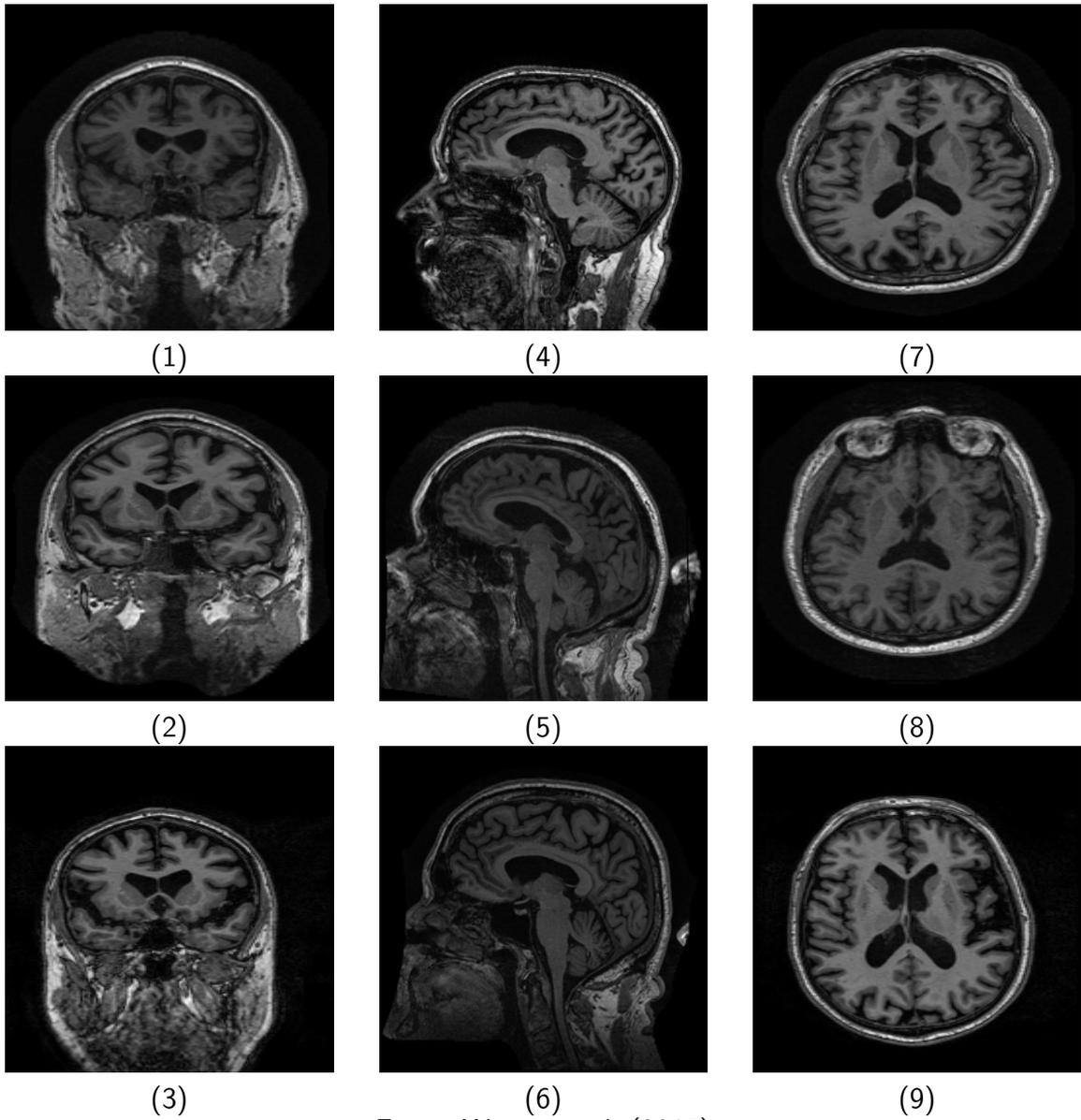
agressão, euforia, desinibição, irritabilidade, labilidade, apatia e comportamento motor aberrante (CUMMINGS, 1997; CUMMINGS et al., 1994).

- Inventário Neuropsiquiátrico Q (NPI-Q): uma versão do NPI para ser administrada de forma mais rápida e por telefone (KAUFER et al., 2000).
- Brain Health Registry (BHR): uma bateria de testes on-line aprovada pelo IRB da Universidade da Califórnia em San Francisco, incluindo perguntas de autorrelato, testes neuropsicológicos e avaliações por parceiros de estudo (ASHFORD et al., 2020). BHR pode ser usado para avaliação na clínica e em casa (ASHFORD et al., 2020).
- Geriatric Depression Scale (GDS) Short Form: uma escala de autorrelato projetada para rastrear sintomas de depressão em idosos (SHEIKH, 1986).

Este trabalho usou 584 volumes de ressonância magnética estrutural ponderados em T1 digitalizados em 3T e extraídos do banco de dados ADNI-3, registrados após 1, 2 e 3 anos de observação de cada paciente (WEINER et al., 2017; YUSHKEVICH et al., 2010; GUNTER et al., 2017). Todos os volumes adquiridos são compostos por $256 \times 256 \times 256$ voxels. Não foi realizado pré-processamento *offline* nas imagens utilizadas neste estudo, pois todas as imagens baixadas já foram submetidas ao pré-processamento online pelos scanners de acordo com o protocolo estabelecido no ADNI-1 que estabelece as seguintes melhorias: correção de gradwarp, correção B1 e correção N3. As imagens baixadas foram excluídas do estudo se apresentassem problemas na definição dos planos do corpo humano ou aquelas cujo conteúdo do arquivo estivesse corrompido. Dois cenários foram investigados neste estudo: com e sem decapagem de crânio/segmentação de extração cerebral automatizada como etapa de pré-processamento online.

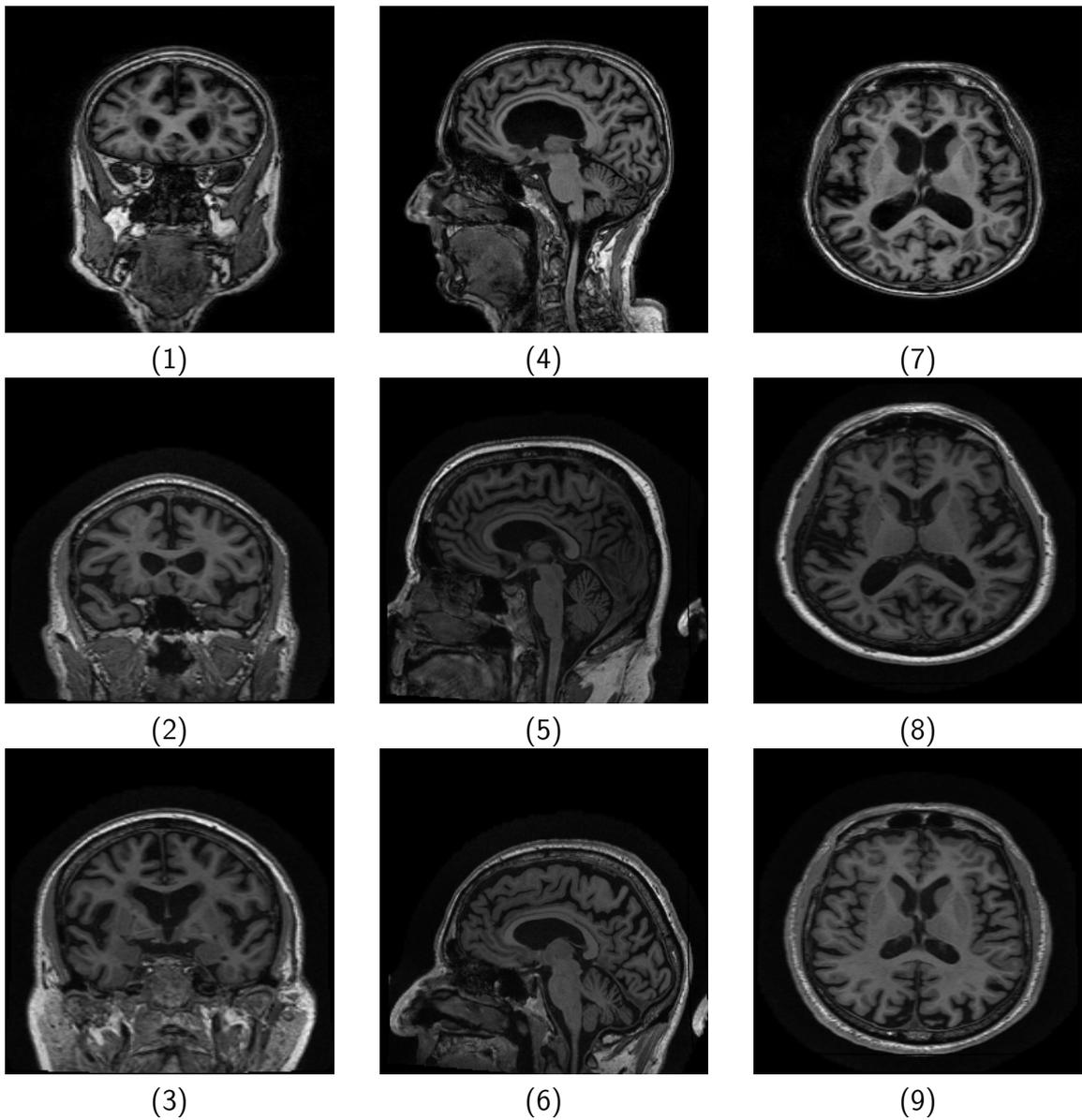
As Figuras 13, 14 e 15 apresentam exemplos de cortes de volume selecionados para CN, MCI e AD para indivíduos do sexo masculino, respectivamente, com cortes coronal, sagital e axial. As Figuras 16, 17 e 18 apresentam exemplos de fatias semelhantes para CN, MCI e AD para mulheres, nesta ordem.

Figura 13 – Exemplos de fatias de volumes ponderados em T1 de indivíduos do sexo masculino cognitivamente normais, para visualizações coronal (1, 2 e 3), sagital (4, 5 e 6) e axial (7, 8 e 9).



Fonte: Weiner et al. (2017)

Figura 14 – Exemplos de fatias de volumes ponderados em T1 para indivíduos do sexo masculino com comprometimento cognitivo leve, para visualizações coronal (1, 2 e 3), sagital (4, 5 e 6) e axial (7, 8 e 9).



Fonte: Weiner et al. (2017)

Figura 15 – Exemplos de fatias de volume ponderados em T1 para indivíduos do sexo masculino com doença de Alzheimer, para visualização coronal (1, 2 e 3), sagital (4, 5 e 6) e axial (7, 8 e 9).

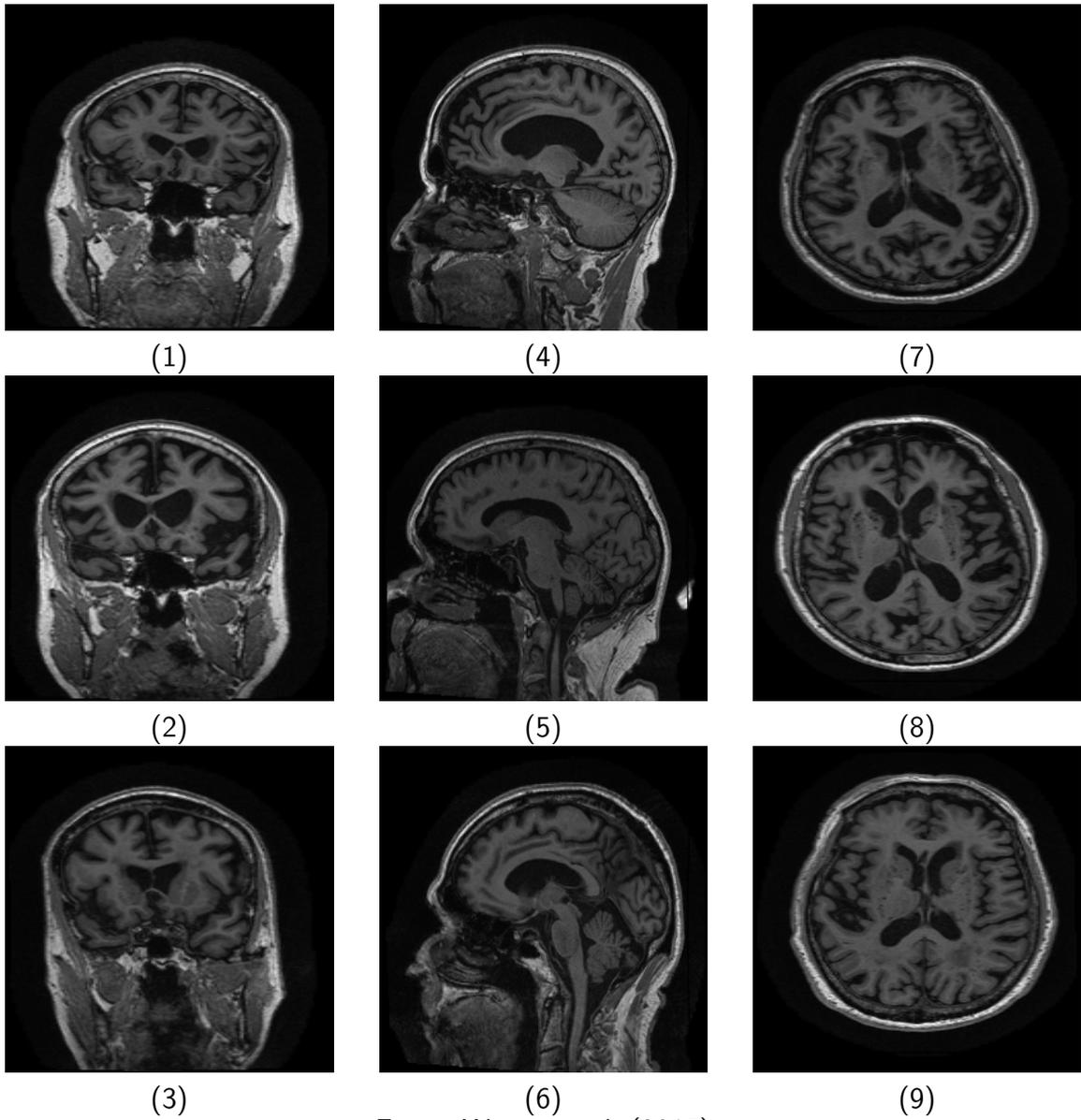
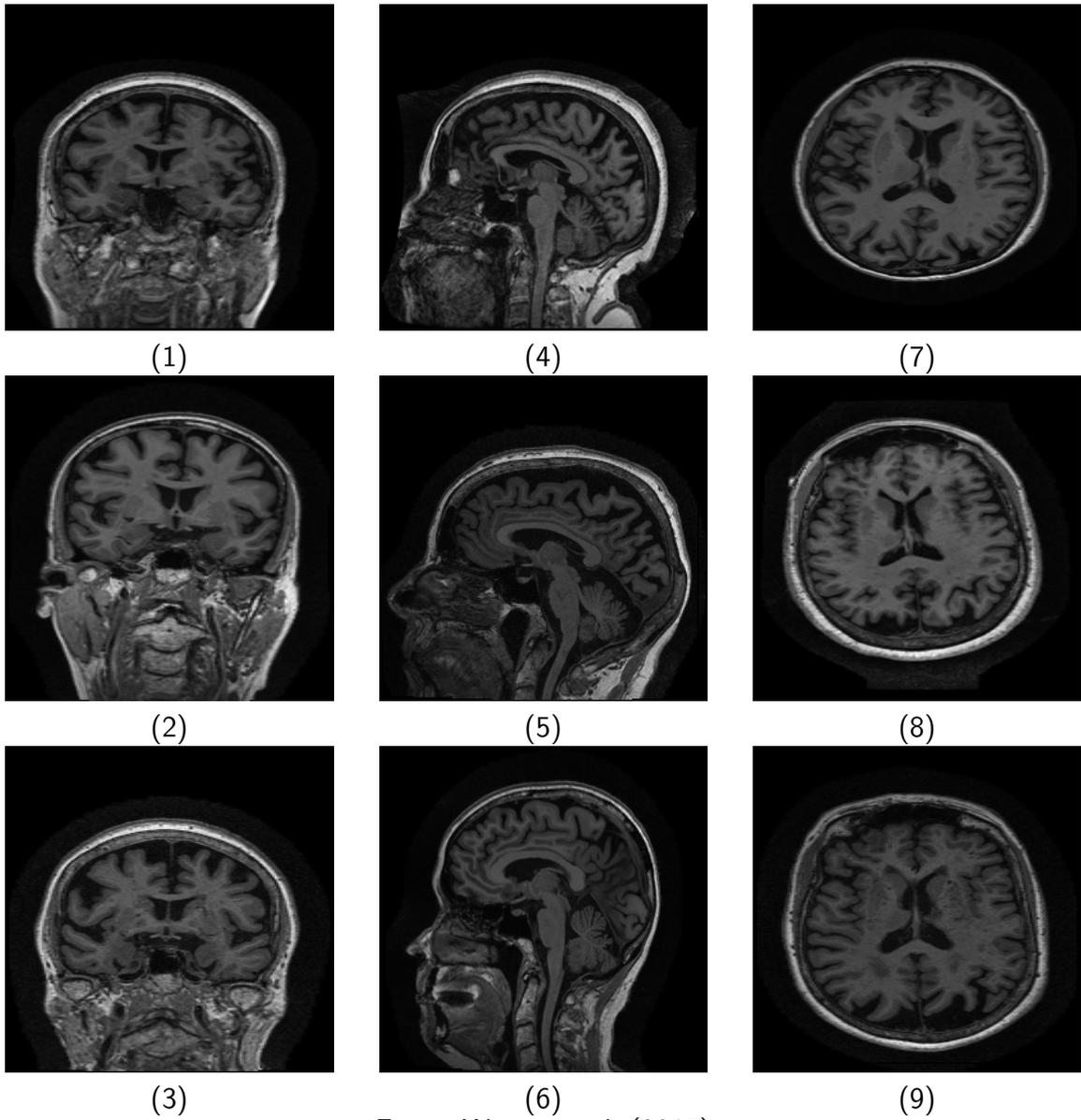
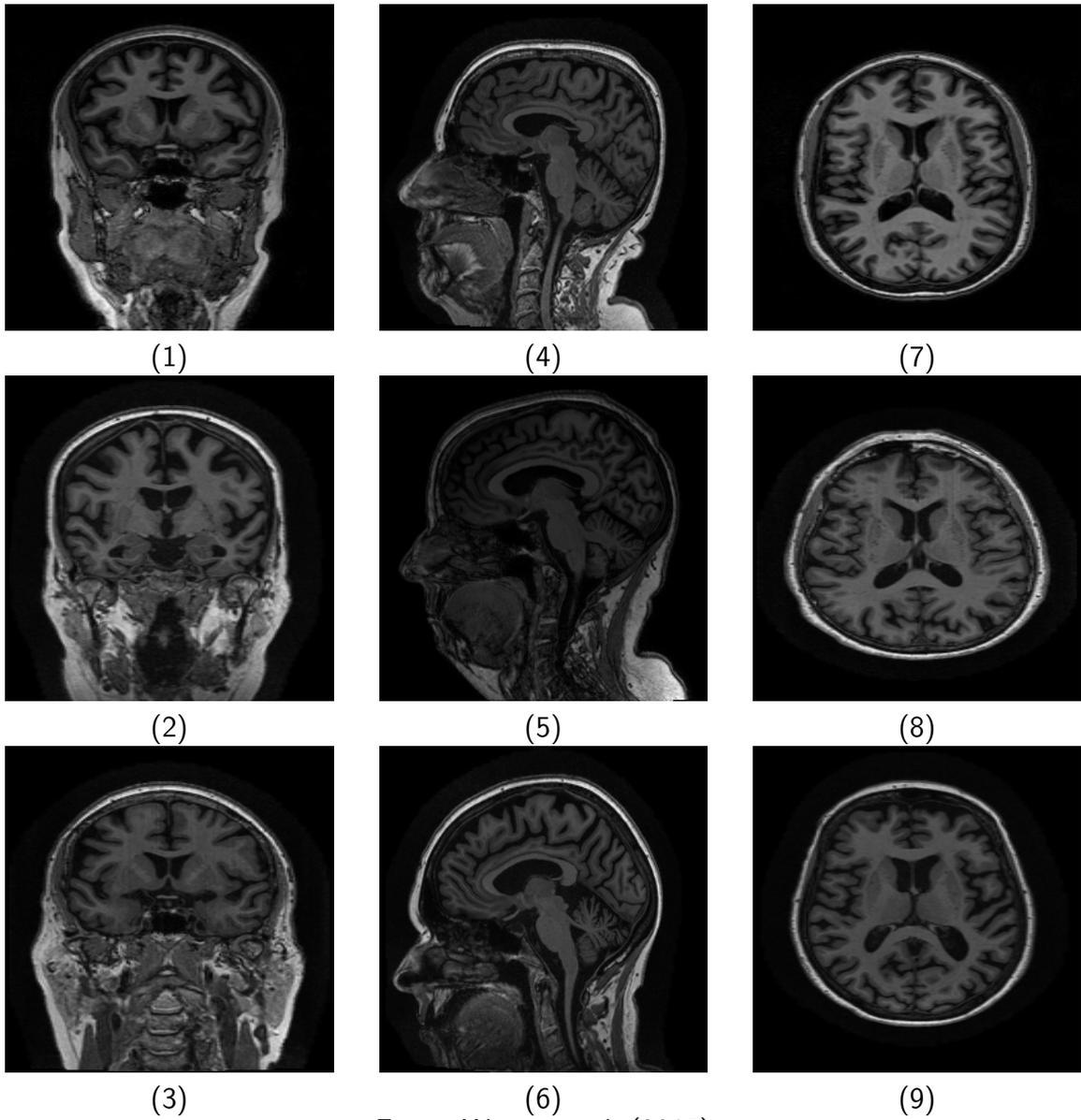


Figura 16 – Exemplos de fatias de volume ponderados em T1 para indivíduos do sexo feminino cognitivamente normais, para visualizações coronal (1, 2 e 3), sagital (4, 5 e 6) e axial (7, 8 e 9).



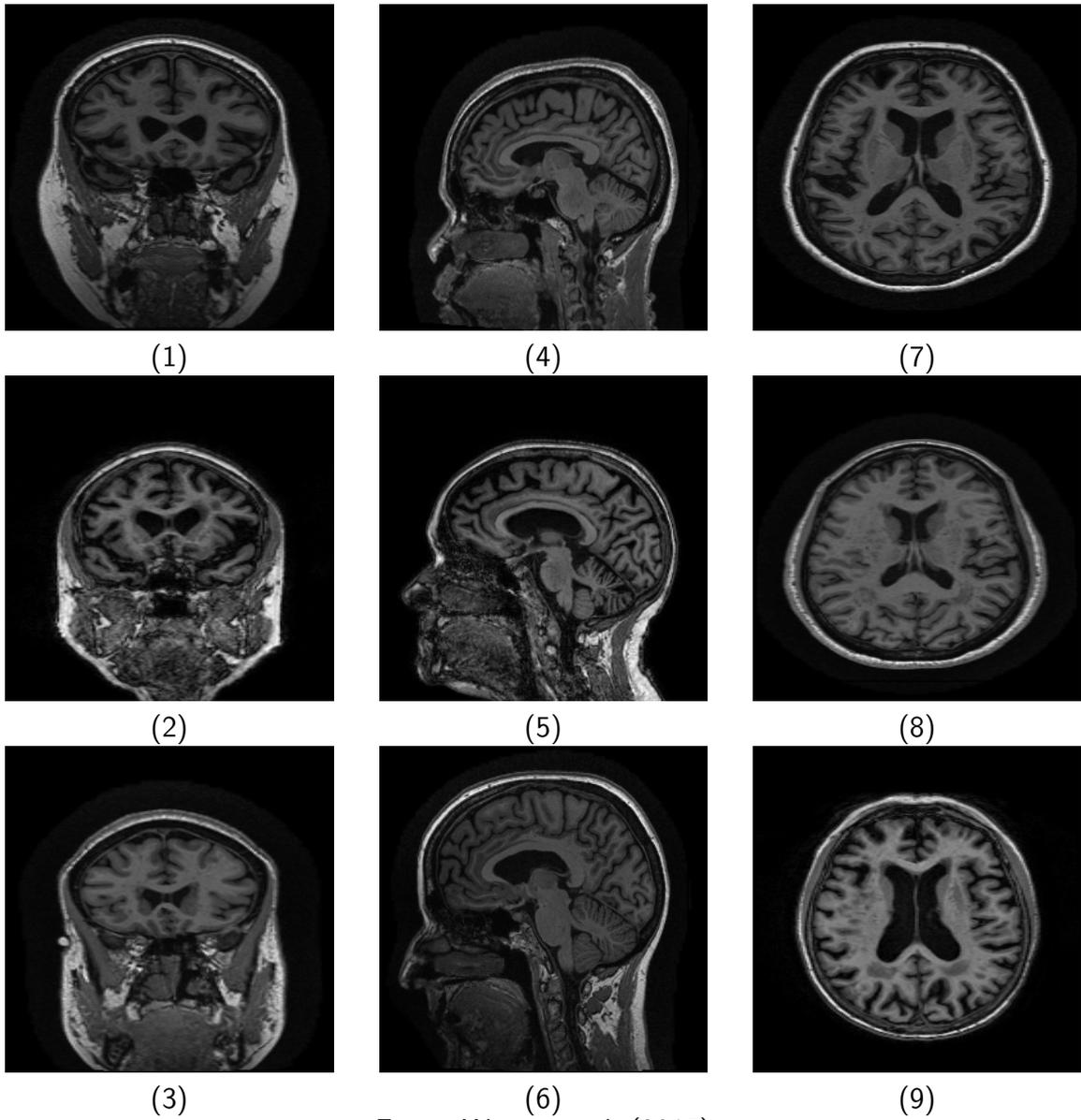
Fonte: Weiner et al. (2017)

Figura 17 – Exemplos de fatias de volume ponderados em T1 para mulheres com comprometimento cognitivo leve, para vistas coronal (1, 2 e 3), sagital (4, 5 e 6) e axial (7, 8 e 9).



Fonte: Weiner et al. (2017)

Figura 18 – Exemplos de fatias de volume ponderados em T1 para mulheres com doença de Alzheimer, para vistas coronal (1, 2 e 3), sagital (4, 5 e 6) e axial (7, 8 e 9).



Fonte: Weiner et al. (2017)

4.8 PRÉ-PROCESSAMENTO: EXTRAÇÃO CEREBRAL AUTOMATIZADA

A segmentação de imagens cerebrais é uma importante etapa de pré-processamento em sistemas para apoiar o diagnóstico de doenças cerebrais. O processo de segmentação cerebral consiste em separar as estruturas cerebrais mais importantes, como cerebelo, córtex, ventrículos e líquido cefalorraquidiano, daquelas que não fazem parte do cérebro, como a caixa craniana, os olhos e outras estruturas.

Considerando volumes de imagens cranianas, a segmentação manual não seria tecnicamente viável. Para realizar essa tarefa, seriam necessários especialistas em neuroimagem, capazes de separar manualmente estruturas como a caixa craniana e os olhos dessas estruturas cerebrais, como o córtex, o cerebelo e os ventrículos. Considerando que todo especialista humano está sujeito a erros por estresse e cansaço, ou mesmo erros pontuais, seriam necessários comitês de especialistas humanos. Considerando também a baixa sensibilidade dos especialistas humanos para detectar estruturas mais detalhadas, os comitês de especialistas humanos em neuroimagem seriam, além de excessivamente caros, não isentos de erros. A segmentação de imagens cerebrais requer uma combinação de análise local e análise global, a fim de considerar tanto aspectos morfológicos gerais quanto detalhes locais, como cantos de imagem e outros detalhes estruturais. Tarefas dessa natureza são caras e tecnicamente inviáveis de serem executadas por especialistas humanos.

Existem pelo menos três abordagens para segmentar imagens cerebrais: (1) métodos manuais; (2) métodos que combinam limiarização e análise morfológica; e (3) métodos adaptativos baseados em superfície (SMITH, 2002).

Na segmentação por limiar e análise morfológica, inicialmente é realizada uma binarização utilizando um limiar inferior e um limiar superior (SMITH, 2002). Esses limiares são usados para separar o tecido cerebral, que geralmente possui níveis de cinza de valor intermediário, das partes mais escuras, como a parte inferior da imagem e a caixa craniana, e das partes mais claras, como os olhos. O resultado é uma imagem binária usada como máscara sobre a imagem original. A análise morfológica é então utilizada para realizar uma segunda segmentação na estrutura remanescente, separando o córtex, cerebelo e ventrículos de outras estruturas não consideradas parte do cérebro. Esses algoritmos são relativamente difíceis de implementar, dada a natureza diferente das imagens obtidas com diferentes sequências de pulso em imagens de ressonância magnética (SMITH, 2002). A definição de limiares para a segmentação inicial também é bastante difícil, de natureza empírica. Assim, esses métodos são considerados, via

de regra, semiautomáticos (SMITH, 2002).

Em métodos adaptativos baseados em superfícies, modelos de superfícies deformáveis são usados (SMITH, 2002). Um modelo de superfície deformável pode ser definido como uma malha tessellada de triângulos. Este modelo é então ajustado à área ocupada pela região do cérebro, ou seja, córtex, cerebelo e ventrículos. Computacionalmente, esse processo requer a imposição de duas restrições principais ao ajuste: (1) uma restrição que impõe uma forma de suavidade na superfície, para manter a superfície bem condicionada e também corresponder à suavidade física da superfície real do cérebro; (2) e uma restrição que adapta o modelo à superfície do cérebro através de deformações iterativas (SMITH, 2002). Em geral, essa abordagem é mais robusta e mais fácil de automatizar com sucesso do que os métodos de limiar com morfologia (SMITH, 2002).

FSL-Brain Extraction Tool (BET) é um método muito popular baseado em superfície adaptativa para segmentação cerebral (VICHIANIN et al., 2018; POPESCU et al., 2012). O algoritmo BET pode ser brevemente descrito como segue. Primeiro, a imagem é binarizada por limiarização, buscando separar a área ocupada pelo cérebro do fundo da imagem e de outras estruturas não cerebrais. Então, o centro de gravidade na imagem binária é calculado. Tendo este centro de gravidade calculado como sua origem, a circunferência circundante de raio máximo é plotada. A partir deste limite circular inicial, a fronteira obtida é ajustada iterativamente através de sucessivas deformações da fronteira e do modelo digital do cérebro. Esta superfície do cérebro é modelada por uma tesselação de superfície usando triângulos conectados. O modelo inicial é uma esfera tessellada, gerada começando com um icosaedro e subdividindo cada triângulo iterativamente em quatro triângulos menores. A distância do vértice é ajustada a partir do centro para formar uma superfície o mais esférica possível. Esta é uma tesselação comum da esfera. Cada vértice tem cinco ou seis vizinhos, de acordo com sua posição em relação ao icosaedro original. O processo iterativo de ajuste de superfície e contorno segue até que o máximo de iterações do DSC mínimo seja alcançado (SMITH, 2002). A métrica DSC é definida da seguinte forma:

$$DSC = \frac{2TP}{2TP + FP + FN}, \quad (4.6)$$

onde DSC é o índice de Sørensen–Dice, um número entre 0 e 1 que indica a fidelidade da imagem binarizada à imagem real (SØRENSEN, 1948; DICE, 1945); TP, FP e FN são os números de pixels verdadeiros positivos, falsos positivos e falsos negativos, respectivamente.

É importante notar que o bom desempenho de segmentação das diferentes versões do BET

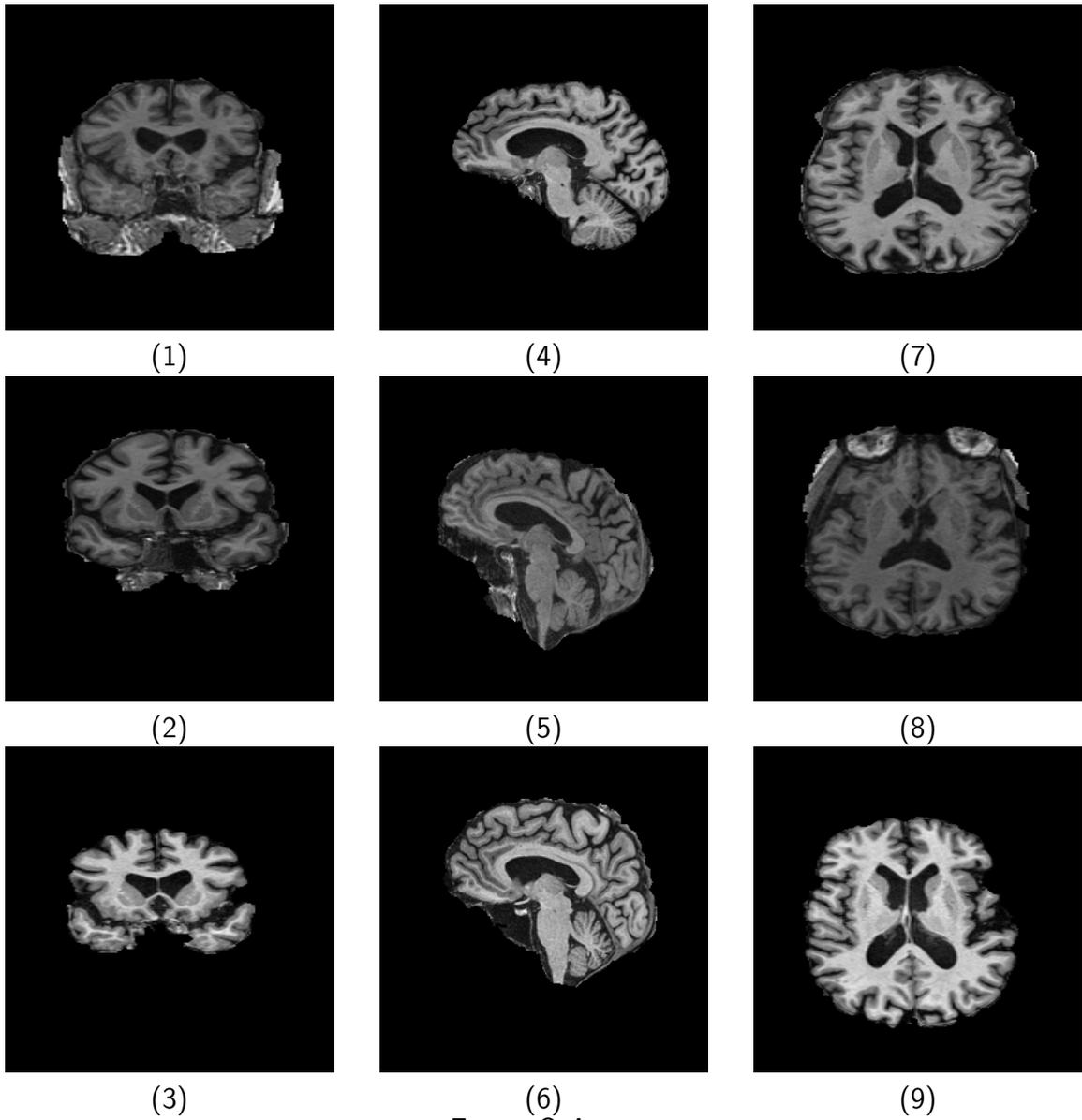
depende muito dos atlas cerebrais utilizados. Segundo Dickie et al. (2017), os atlas têm importantes aplicações em pesquisas como alvos de registro e segmentação em estudos de imagem populacional e na prática clínica. Os modelos são úteis para identificar alterações estruturais cerebrais fora dos limites normais e, cada vez mais, para uso no planejamento cirúrgico. Dickie et al. (2017) realizou uma revisão sobre atlas cerebrais, apontando os principais problemas e limitações dos mais importantes atlas disponíveis, principalmente a falta de informações clínicas e demográficas úteis para caracterizar volumes cerebrais em alguns volumes de imagem.

O algoritmo BET passou por várias modificações, atualizações e melhorias (POPESCU et al., 2012; VICHIANIN et al., 2018; ISENSEE et al., 2019; SHI et al., 2012; LEUNG et al., 2011; WANG et al., 2021; HECKEMANN et al., 2015; EZHILARASAN et al., 2021; CHUNG; PARK, 2020). Vichianin et al. (2018) busca por parâmetros ideais relacionados à extração cerebral automatizada em ressonância magnética usando ferramentas BET e FreeSurfer. A comparação é feita com o padrão ouro que corresponde à extração feita manualmente. Este estudo retrospectivo foi realizado em 10 indivíduos adultos saudáveis com dados de 3D-T1W em máquina de RM 3T. As medições manuais de volume cerebral padrão ouro foram feitas por dois leitores independentes. As segmentações automatizadas utilizando BET com parâmetros variados e software FreeSurfer também foram realizadas. Em seguida, os dois métodos automatizados foram comparados com o traçado manual para fazer o parâmetro ótimo buscando o maior coeficiente de correlação intraclass (ICC). De acordo com Vichianin et al. (2018), em diferentes situações, pesquisas com segmentação automatizada precisam ser ajustadas e validadas com um padrão ouro aceitável antes da implementação clínica. Popescu et al. (2012) realizaram um trabalho semelhante em relação à otimização de parâmetros BET. Seu objetivo foi analisar sistematicamente o desempenho do BET em imagens de pacientes com esclerose múltipla, variando seus parâmetros e combinações de opções, e comparando quantitativamente seus resultados com um padrão ouro manual.

Neste trabalho, foi utilizada a versão BET2 implementada usando a biblioteca Python Nipype de acordo com a definição de Jenkinson et al. (2005).

As Figuras 19, 20 e 21 apresentam exemplos de fatias de volume selecionadas após BET para CN, MCI e AD para indivíduos do sexo masculino, respectivamente, nas visualizações axial, sagital e coronal. As Figuras 22, 23 e 24 apresentam exemplos de fatias semelhantes após BET para CN, MCI e AD para mulheres, nesta ordem.

Figura 19 – Exemplos de fatias de volumes ponderados em T1 após BET para indivíduos do sexo masculino cognitivamente normais, para visualizações coronal (1, 2 e 3), sagital (4, 5 e 6) e axial (7, 8 e 9).



Fonte: O Autor

Figura 20 – Exemplos de fatias de volumes ponderados em T1 após BET para indivíduos do sexo masculino com comprometimento cognitivo leve, para visualizações coronal (1, 2 e 3), sagital (4, 5 e 6) e axial (7, 8 e 9) .

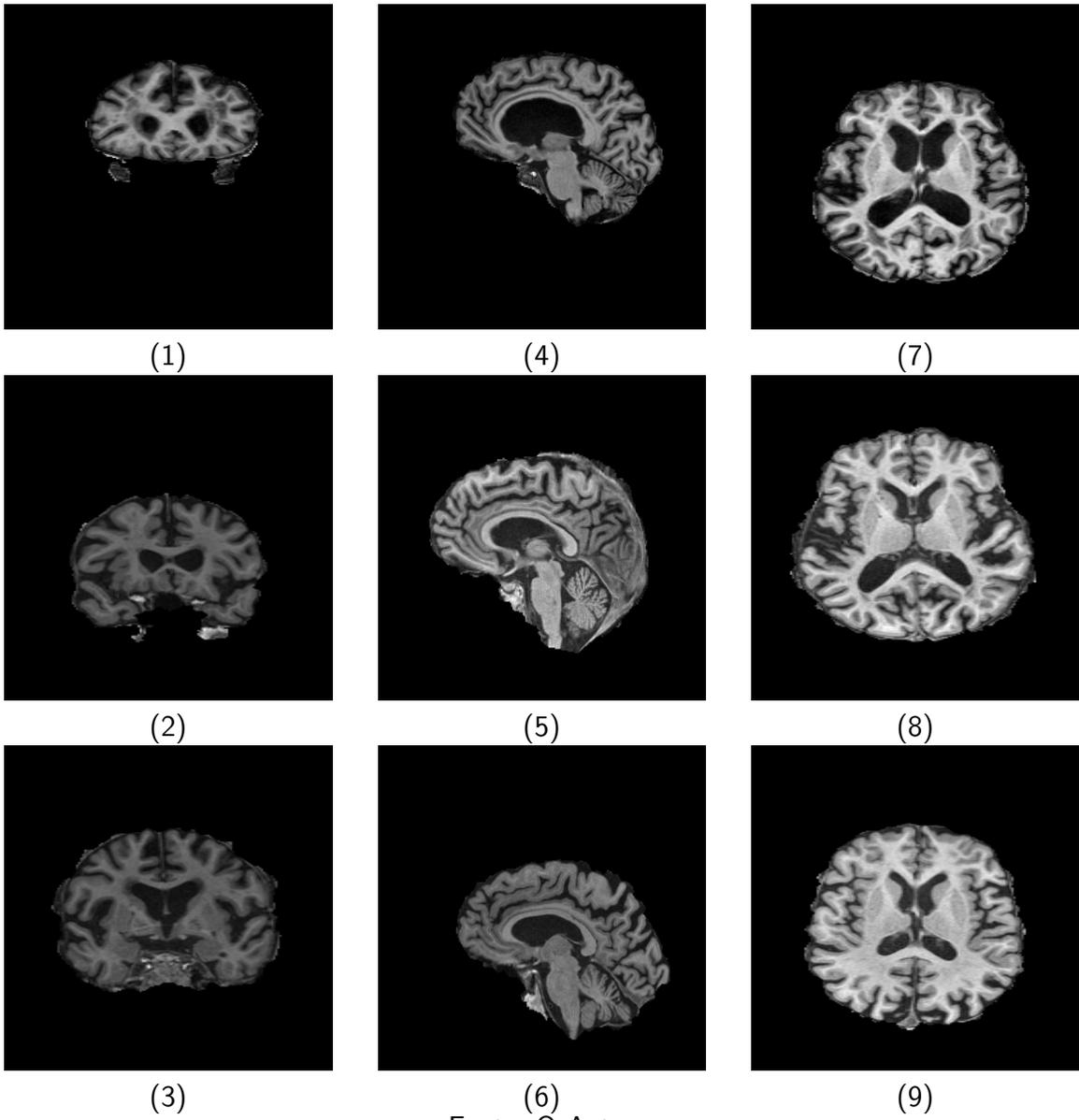


Figura 21 – Exemplos de fatias de volumes ponderados em T1 após BET para indivíduos do sexo masculino com doença de Alzheimer, para vistas coronal (1, 2 e 3), sagital (4, 5 e 6) e axial (7, 8 e 9).

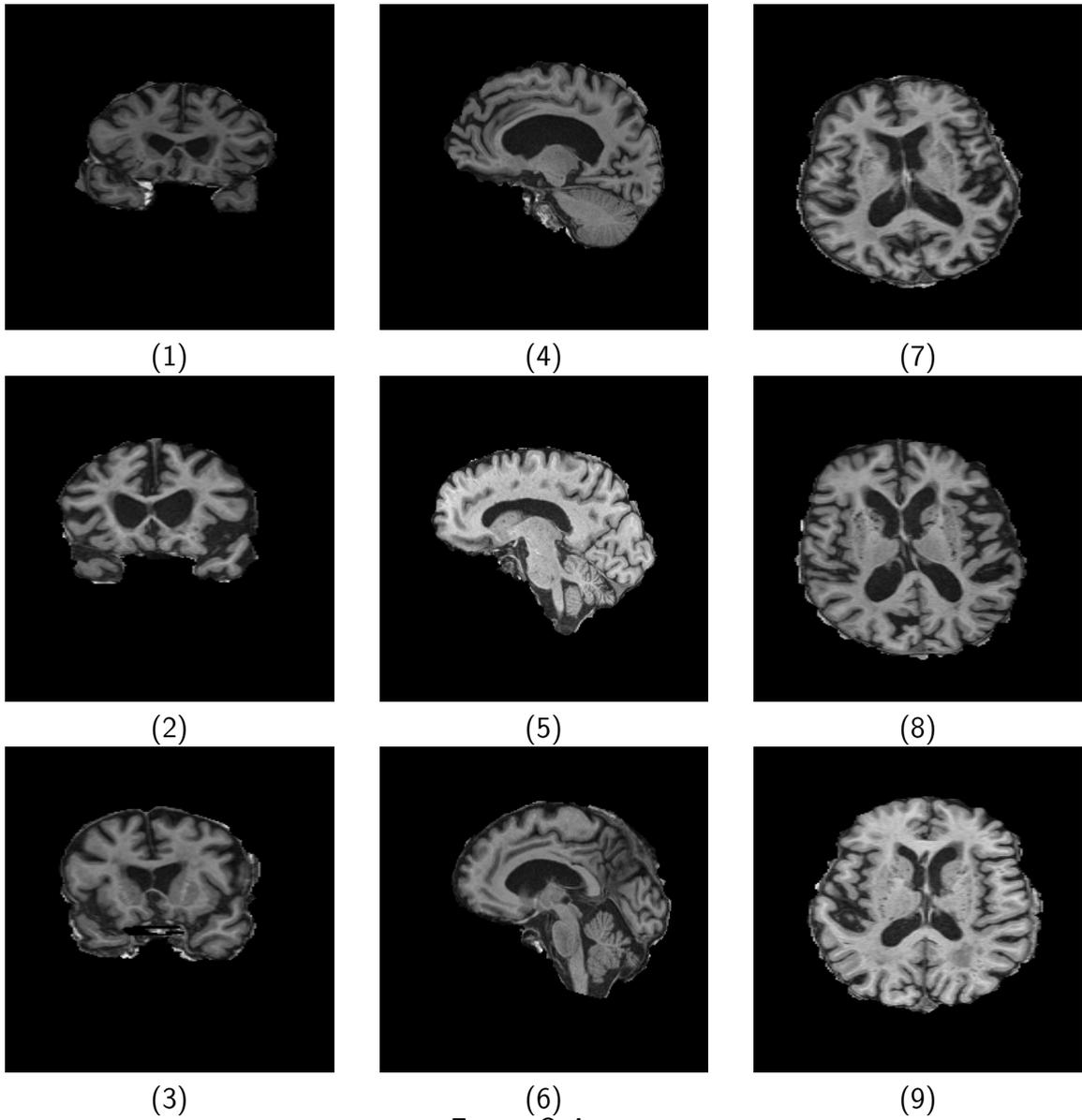
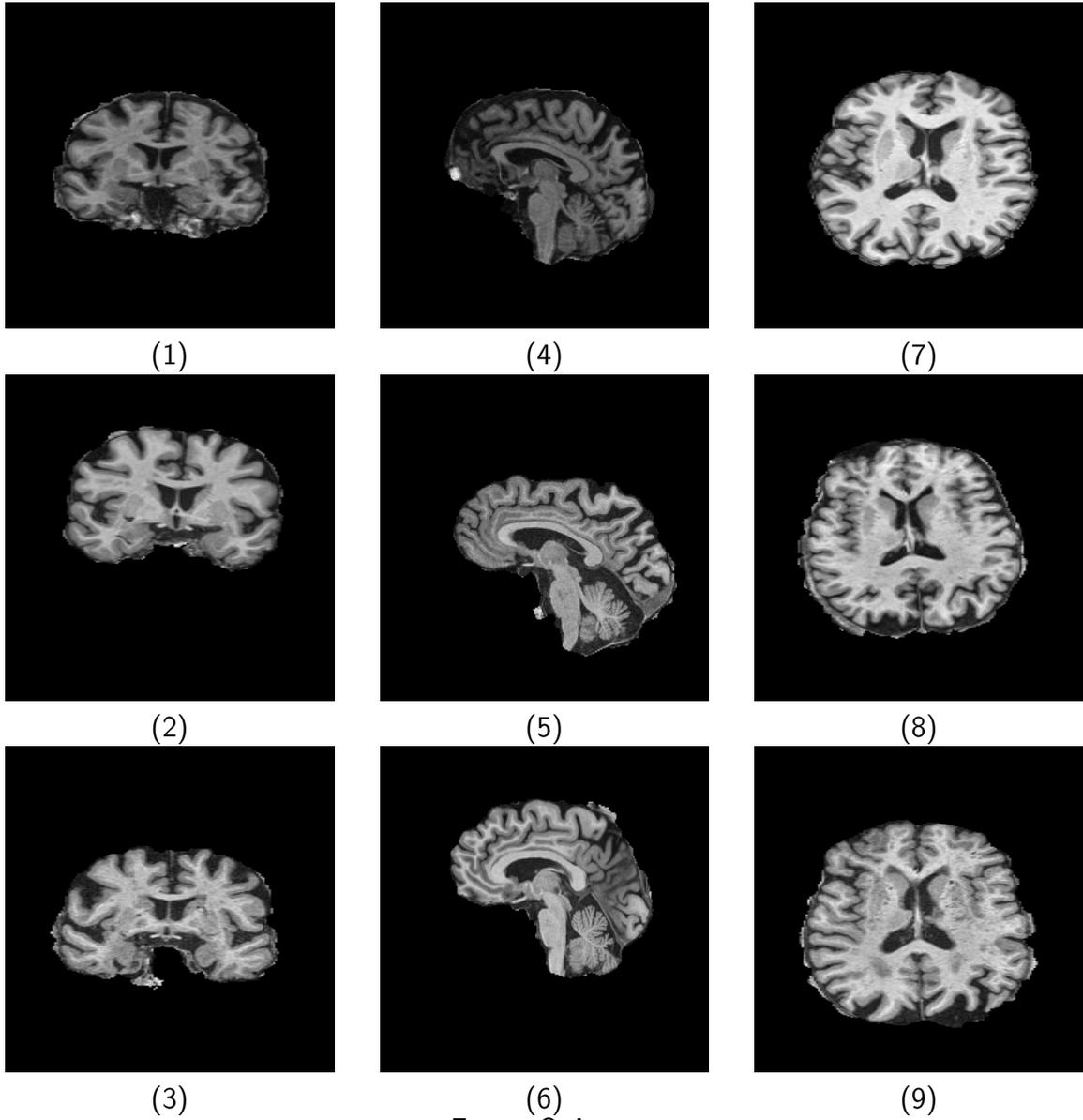
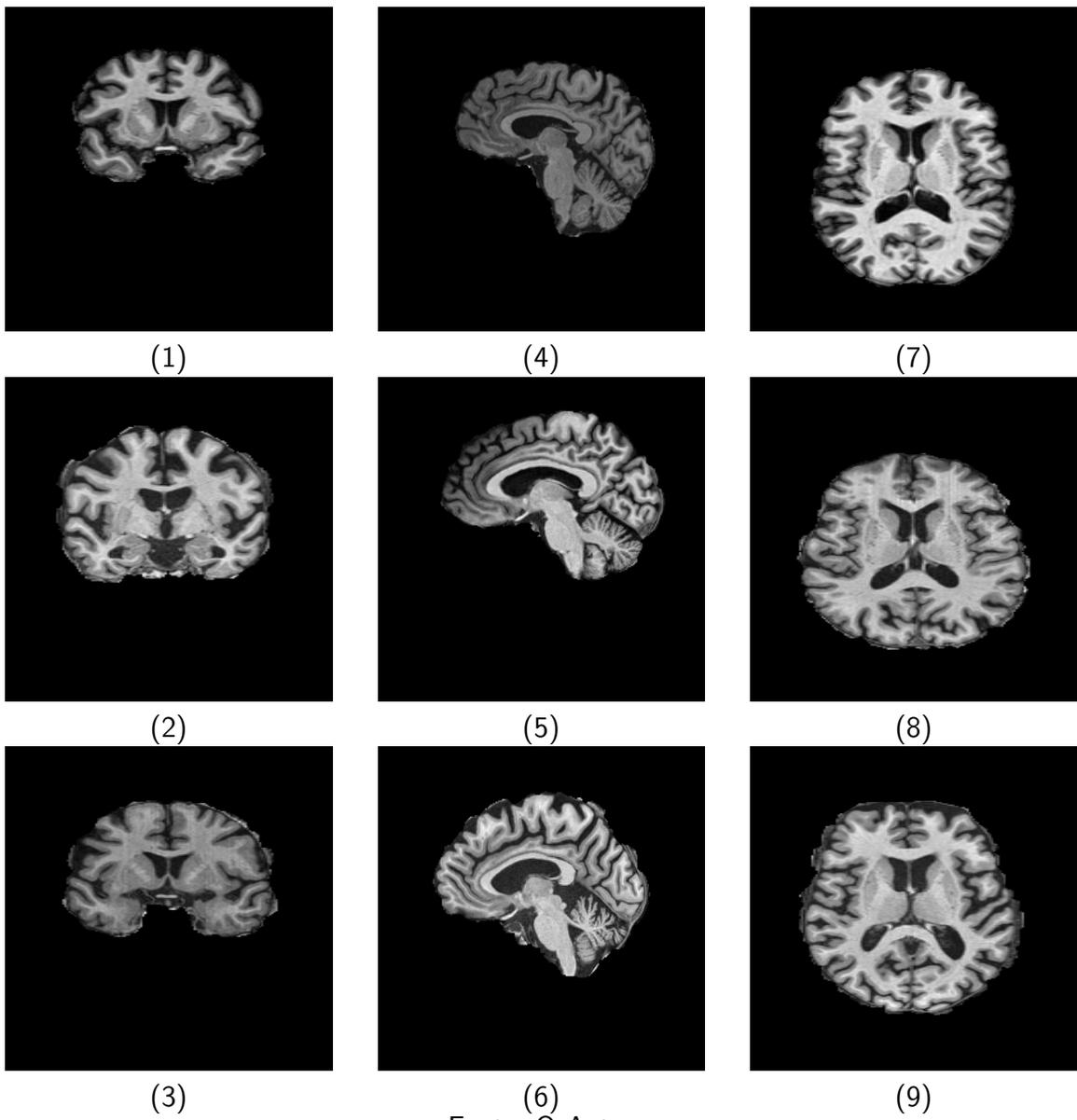


Figura 22 – Exemplos de fatias de volumes ponderados em T1 após BET para mulheres cognitivamente normais, para vistas coronal (1, 2 e 3), sagital (4, 5 e 6) e axial (7, 8 e 9).



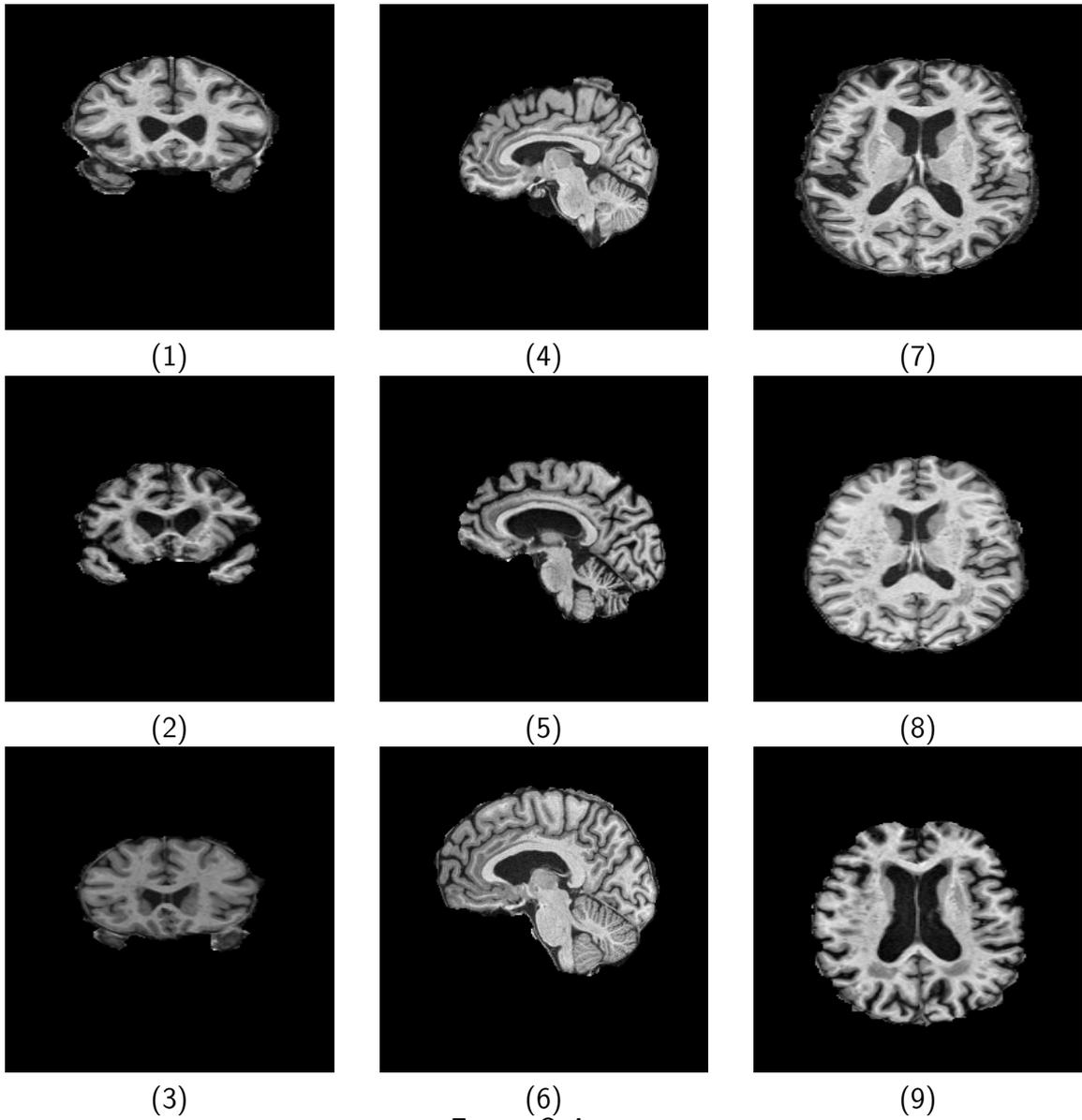
Fonte: O Autor

Figura 23 – Exemplos de fatias de volumes ponderados em T1 após BET para indivíduos do sexo feminino com comprometimento cognitivo leve, para visualizações coronal (1, 2 e 3), sagital (4, 5 e 6) e axial (7, 8 e 9).



Fonte: O Autor

Figura 24 – Exemplos de fatias de volumes ponderados em T1 após BET para mulheres com doença de Alzheimer, para vistas coronal (1, 2 e 3), sagital (4, 5 e 6) e axial (7, 8 e 9).



4.9 BALANCEAMENTO DE CLASSES MINORITÁRIA POR MEIO DO MÉTODO SMOTE

Chawla et al. (2002) propôs uma abordagem de superamostragem na qual a classe minoritária em um conjunto de dados desbalanceado é superamostrada criando exemplos sintéticos em vez de superamostrar esse conjunto de dados com substituição. Essa abordagem é inspirada em uma técnica que provou ser bem-sucedida no reconhecimento de caracteres manuscritos: dados de treinamento extras são criados executando certas operações para perturbar os dados de treinamento. Chawla et al. (2002) propõe a geração de exemplos sintéticos operando em espaço de características. A classe minoritária é superamostrada tomando cada amostra de classe minoritária e introduzindo exemplos sintéticos ao longo dos segmentos de linha que unem todos os k vizinhos mais próximos da classe minoritária. Os vizinhos dos k vizinhos mais próximos são escolhidos aleatoriamente. Na proposta original (CHAWLA et al., 2002), as amostras sintéticas são geradas da seguinte forma: a) Pegue a diferença entre o vetor de características (amostra) em consideração e seu vizinho mais próximo. b) Multiplique essa diferença por um número aleatório entre 0 e 1 e adicione-o ao vetor de características em consideração. Isso causa a seleção de um ponto aleatório ao longo do segmento de linha entre dois recursos específicos. Essa abordagem efetivamente força a região de decisão da classe minoritária a se tornar mais geral. Vários trabalhos melhoraram esse método adicionando procedimentos ligeiramente diferentes para gerar vetores de recursos sintéticos por um conjunto de k vizinhos aleatórios (HAN; WANG; MAO, 2005; FERNÁNDEZ et al., 2018; TORGO et al., 2013; BUNKHUM-PORNPAT; SINAPIROMSARAN; LURSINSAP, 2009; MACIEJEWSKI; STEFANOWSKI, 2011; DOUZAS; BACAO, 2019; DOUZAS; BACAO; LAST, 2018).

Dado um conjunto de m classes, $\Omega = \{\omega_1, \omega_2, \dots, \omega_m\}$ e um universo de n -dimension de recursos preenchidos com vetores de recursos $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$, considerando uma classe minoritária ω_j com instâncias m_j , onde $j = 1, 2, \dots, m$; considerando que cada classe ω_j com instâncias m_j precisa de mais instâncias M_j para balancear o conjunto de treinamento, ou seja, $m_i + M_i = m_j + M_j$, e $i \neq j, \forall i, j \in \{1, 2, \dots, m\}$; para cada classe minoritária ω_j , uma instância é escolhida aleatoriamente e, em seguida, as instâncias vizinhas $(k - 1)$ mais próximas também são selecionadas, dando um conjunto de k instâncias $\{\mathbf{x}_p\}_{p=1}^k$; então, o conjunto de instâncias sintéticas $\{\mathbf{x}'_q\}_{q=1}^{N_j}$ da seguinte forma:

$$\mathbf{x}'_q = \mathbf{x}_u + \frac{\sum_{p=1}^k r_p (\mathbf{x}_p - \mathbf{x}_u)}{\sum_{p=1}^k r_p}, \quad (4.7)$$

onde $r_p \sim R(0, 1)$ é um número real aleatório distribuído de acordo com uma função de

distribuição probabilística arbitrária $R : \mathbb{R} \rightarrow [0, 1]$; u é um índice aleatório do conjunto de vizinhos mais próximos, $1 \leq u \leq k$; e $q = 1, 2, \dots, N_j$.

4.10 ARRANJO EXPERIMENTAL

Todos os experimentos foram realizados usando algoritmos de aprendizado de máquina configurados para classificação multiclasse que fornecem um valor de acurácia específico para cada classe: Cognitivamente Normal (CN), Comprometimento Cognitivo Leve (MCI) e Doença de Alzheimer (DA). No entanto, uma abordagem “um contra todos” foi usada para nos permitir calcular as principais métricas de diagnóstico como especificidade, sensibilidade, classe vs all acurácia, F1-score e precisão. Portanto, os resultados obtidos nos experimentos são agrupados na forma de três classificações binárias e uma classificação multiclasse da seguinte forma:

- CN vs MCI vs AD: Classificação Multiclasse;
- CN vs Todas as classes: classificação binária;
- MCI vs Todas as classes: classificação binária;
- AD vs Todas as classes: classificação binária.

Para realização dos experimentos, foram utilizadas as implementações existentes na biblioteca scikit-learn dos classificadores empregados em nossos experimentos. Adotamos o scikit-learn 0.20.3 com o Python 3.6.8. Para aplicar o algoritmo genético com representação inteira, foi adotada a biblioteca Distributed Evolutionary Algorithms in Python - DEAP. Para calcular as estatísticas de Haralick (HARALICK; DINSTEIN; SHANMUGAM, 1973) e dos Momentos de Zernike (KHOTANZAD; HONG, 1990), foi utilizada a versão 1.4.5 da biblioteca Python Mahotas (COELHO, 2013).

5 RESULTADOS E DISCUSSÃO

Neste capítulo são apresentados os resultados quantitativos e qualitativos e descritos os arranjos experimentais montados para a obtenção desses resultados. São também feitas discussões e tecidas conclusões parciais sobre a metodologia proposta.

5.1 RESULTADOS EXPERIMENTAIS

Neste trabalho, os experimentos foram realizados considerando dois cenários: com e sem segmentação BET como etapa de pré-processamento da imagem. Em cada um desses cenários foram considerados três grupos de imagens com base no gênero dos sujeitos da seguinte forma: primeiro grupo com todas as imagens, um segundo apenas com imagens masculinas e um terceiro grupo com apenas imagens femininas. Resumos dos resultados obtidos pelos modelos de classificação propostos, organizados por gênero, são apresentados nas Tabelas 2 e 3, com e sem BET como etapa de pré-processamento, respectivamente. Os resultados são expressos como acurácia geral, para classificação multiclasse, e como especificidade, sensibilidade, precisão, pontuação F1 e acurácia, para os cenários de classificação binária (um contra todos).

Tabela 2 – Resultados agrupados por gênero, usando o pré-processamento BET

Class	Metric	Gender		
		Male	Female	Both
Multiclass classification	global Accuracy	84.17%	83.59%	77.89%
CN vs All	specificity	88.79%	91.93%	87.92%
	sensitivity	93.10%	94.25%	87.44%
	accuracy	89.35%	90.97%	85.71%
	f1-score	91.19%	91.57%	85.96%
	precision	75.25%	81.05%	77.40%
MCI vs All	specificity	97.50%	98.33%	94.58%
	sensitivity	83.16%	83.16%	76.45%
	accuracy	87.42%	87.74%	84.86%
	f1-score	84.96%	85.39%	79.98%
	precision	98.06%	98.71%	93.16%
AD vs All	specificity	96.03%	96.39%	92.28%
	sensitivity	90.91%	93.94%	83.33%
	accuracy	93.87%	94.52%	88.95%
	f1-score	91.42%	93.26%	85.41%
	precision	68.57%	72.22%	64.15%

Fonte: Souza et al. (2021)

Tabela 3 – Resultados agrupados por gênero, sem uso de BET

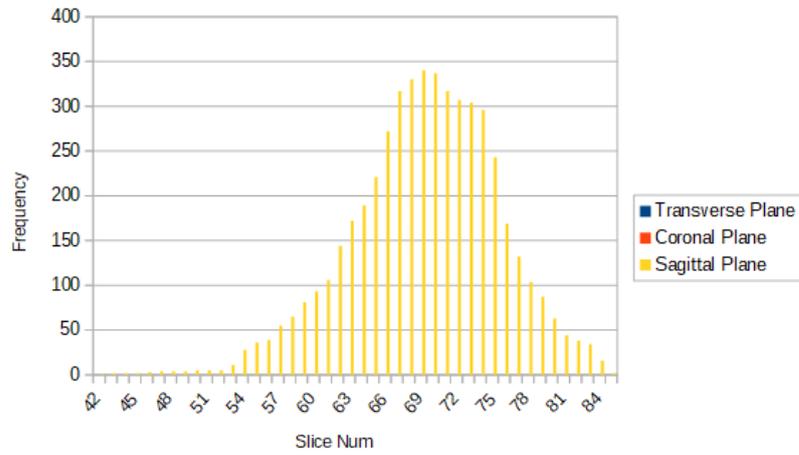
Class	Metric	Gender		
		Male	Female	Both
Muticlass classification	global Accuracy	82.71%	82.02%	78.99%
CN vs All	specificity	91.93%	94.58%	88.69%
	sensivity	93.42%	81.08%	85.56%
	accuracy	91.30%	86.64%	86.46%
	f1-score	92.01%	83.77%	85.75%
	precision	78.57%	89.89%	77.39%
MCI vs All	specificity	98.17%	94.83%	96.47%
	sensivity	80.53%	89.32%	76.11%
	accuracy	86.29%	89.89%	84.90%
	f1-score	83.15%	89.24%	80.03%
	precision	98.64%	90.32%	95.48%
AD vs All	specificity	95.86%	92.99%	90.42%
	sensivity	90.91%	92.06%	87.50%
	accuracy	94.31%	90.61%	89.06%
	f1-score	91.27%	91.33%	87.59%
	precision	71.05%	76.92%	62.70%

Fonte: Souza et al. (2021)

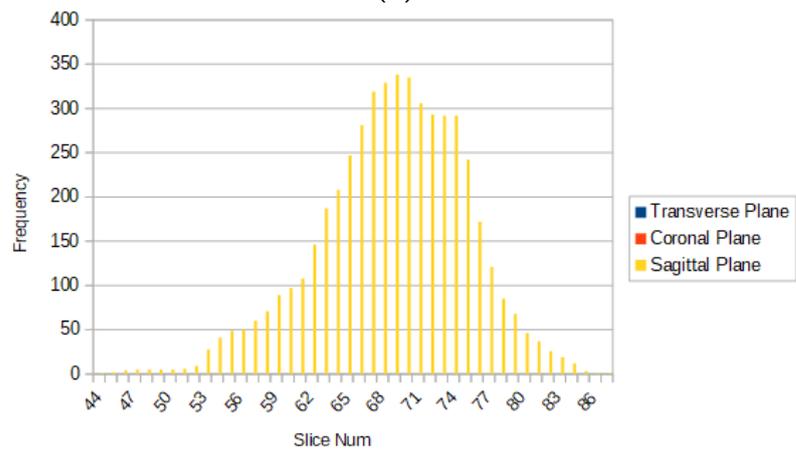
O algoritmo evolutivo foi executado 50 vezes para cada um dos sete algoritmos de classificação e o agrupamento de fatias que obteve a máxima precisão multiclasse foi anotado. Assim, foram realizadas 350 buscas pelo melhor agrupamento de fatias. As Tabelas 4 e 5 apresentam as fatias mais frequentes e sua frequência absoluta e relativa, organizadas de acordo com o sexo (F para feminino, M para masculino e ambos) e plano corporal, para pré-processamento com e sem segmentação BET, respectivamente. As Figuras 25 e 26 exibem histogramas de distribuição de fatias por gênero, com e sem pré-processamento de segmentação BET, respectivamente.

A Tabela 6 mostra os resultados do teste de normalidade Kolmogorov-Smirnov das distribuições de fatias de acordo com o sexo (masculino, feminino ou ambos) e utilizando a segmentação BET como pré-processamento. A Tabela 6 apresenta os resultados da hipótese de Kruskal-Wallis para amostras k-independentes. A Tabela 8 apresenta os resultados dos testes de hipóteses bicaudais de Mann-Whitney Wilcoxon. Seu objetivo é fazer comparações entre resultados para o mesmo gênero e para um determinado gênero contra ambos os gêneros, com ou sem aplicação de BET.

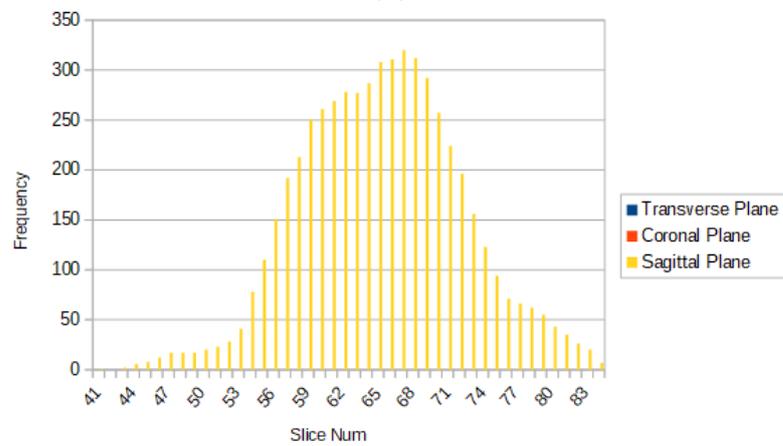
Figura 25 – Histogramas de distribuição de fatias para participantes do sexo feminino (1), masculino (2) e de ambos os sexos (3), considerando o pré-processamento da segmentação BET.



(1)



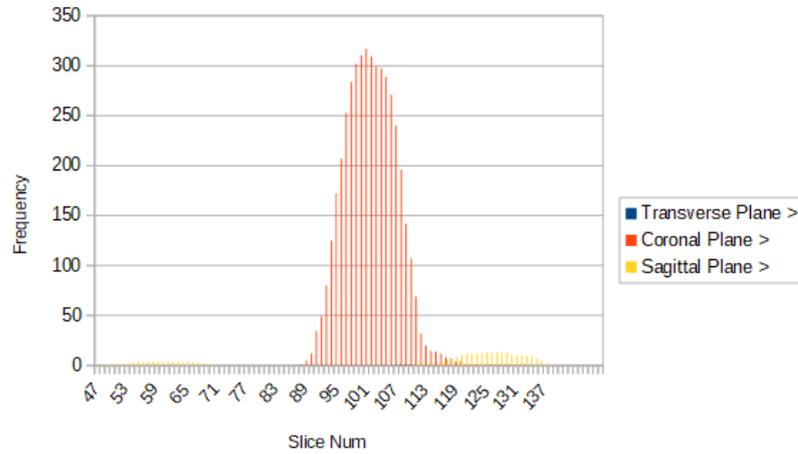
(2)



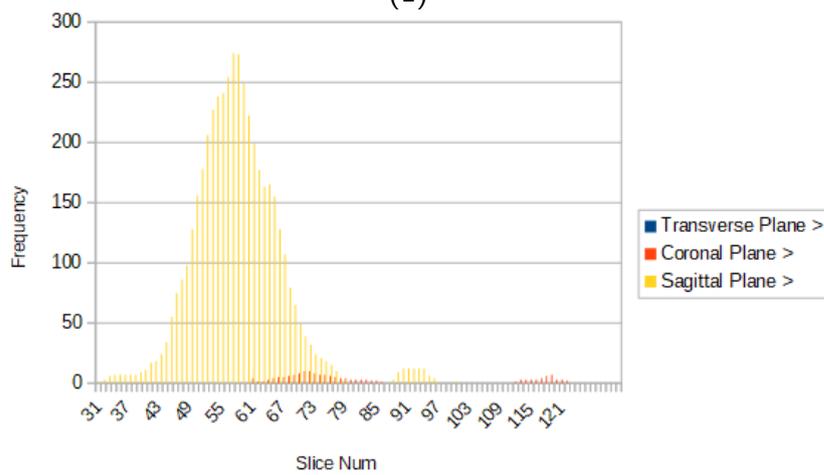
(3)

Fonte: Souza et al. (2021)

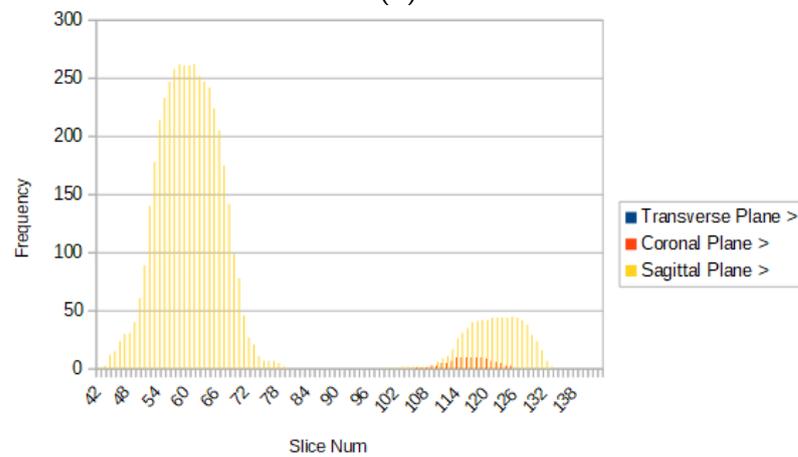
Figura 26 – Histogramas de distribuição de fatias para participantes do sexo feminino (1), masculino (2) e de ambos os sexos (3), sem segmentação BET.



(1)



(2)



(3)

Fonte: Souza et al. (2021)

Tabela 4 – Cortes mais frequentes e frequência absoluta e relativa, segundo sexo (F para feminino, M para masculino e ambos) e plano corporal com pré-processamento BET

Most frequent slices in experiments using BET						
Gender	Body plane	Most frequent slices			Subtotal (all planes)	Total
		Slice	Freq	%		
F	Axial	-	0	0.00%	0	5030
	Coronal	-	0	0.00%	0	
	Sagittal	67	317	6.30%	5030	
		68	330	6.56%		
		69	340	6.76%		
		70	337	6.70%		
71	317	6.30%				
M	Axial	-	0	0.00%	0	5041
	Coronal	-	0	0.00%	0	
	Sagittal	67	319	6.33%	5041	
		68	330	6.55%		
		69	340	6.74%		
		70	337	6.69%		
71	317	6.29%				
Both	Axial	-	0	0.00%	0	5038
	Coronal	-	0	0.00%	0	
	Sagittal	65	308	6.12%	5038	
		66	311	6.18%		
		67	320	6.36%		
		68	312	6.20%		
69	292	5.81%				

Fonte: Souza et al. (2021)

Tabela 5 – Cortes mais frequentes em frequência absoluta e relativa, segundo sexo (F para feminino, M para masculino e ambos) e plano corporal sem pré-processamento BET

Most frequent slices in experiments without BET						
Gender	Body plane	Most frequent slices			Subtotal (all planes)	Total
		Slice	Freq	%		
F	Axial	108	1	0.02%	3	4629
		109	1	0.02%		
		110	1	0.02%		
	Coronal	100	310	6.70%	4490	
		101	317	6.85%		
		102	309	6.68%		
	Sagittal	126	13	0.28%	136	
		127	13	0.28%		
		128	13	0.28%		
M	Axial	-	0	0.00%	0	4774
	Coronal	70	8	0.17%	122	
		71	10	0.22%		
		72	10	0.22%		
		73	8	0.17%		
	Sagittal	56	254	5.49%	4652	
		57	274	5.92%		
		58	273	5.90%		
		59	249	5.38%		
Both	Axial	-	0	0.00%	0	4679
	Coronal	10	115	2.46%	66	
		10	116	2.51%		
		10	117	2.53%		
		10	118	2.55%		
	Sagittal	58	262	5.66%	4613	
		59	261	5.64%		
		60	261	5.64%		
		61	262	5.66%		

Fonte: Souza et al. (2021)

Tabela 6 – Resultados do teste de normalidade Kolmogorov-Smirnov das distribuições de fatias de acordo com o sexo, com e sem segmentação por BET.

Gender	BET	One-Sample Kolmogorov-Smirnov Test (Normalization study)	
		KS value	p value
F	no	1.13	.139
	yes	1.43	.023
M	no	1.48	.017
	yes	1.27	.061
MF	no	.285	.990
	yes	.090	1.21

Fonte: Souza et al. (2021)

Tabela 7 – Resultados da hipótese de Kruskal-Wallis para k amostras independentes

K Samples Kruskal-Wallis Test		K Samples Kruskal-Wallis Test: Ranks Table			
Groups	F, M, MF, F*, M*, MF*	Gender	BET	N	Median Rank
Degrees of freedom	5	F	no	350	932.84
Chi-square	1831.63		yes	350	1711.07
p value	.000	M	no	350	1226.54
			yes	350	1727.82
*BET		MF	no	350	420.89
			yes	350	283.84

Fonte: Souza et al. (2021)

Tabela 8 – Resultados dos testes de hipóteses bicaudais de Mann-Whitney Wilcoxon.

Mann-Whitney two-tailed					
sample I	sample II	Mann-Whitney U	Wilcoxon W	Z value	p value
F	F*	1455.00	62880.00	-22.35	.000
F	MF	1259.00	62684.00	-22.43	.000
F*	MF*	.00	61425.00	-22.90	.000
M	M*	8646.00	70071.00	-19.66	.000
M	MF	3.00	61428.00	-22.90	.000
M*	MF*	.00	61425.00	-22.90	.000
MF	MF*	37876.00	99301.00	-8.74	.000

*BET applied to images

Fonte: Souza et al. (2021)

5.2 DISCUSSÃO

A Tabela 6 mostra o resultado do teste de normalidade Kolmogorov-Smirnov de uma amostra para as distribuições de fatias de acordo com o sexo (masculino, feminino ou ambos) e com ou sem segmentação BET como pré-processamento. Os resultados indicam que, para todas as configurações de gênero e independente do uso de BET no pré-processamento, as distribuições não são normais. Conseqüentemente, não podemos usar o teste de hipótese ANOVA ou qualquer outra métrica paramétrica para comparar os resultados. O teste de Kruskal-Wallis para k amostras independentes revelou, para $p < 0,05$, que pelo menos uma das seis distribuições analisadas (três grupos de gênero sem aplicação de BET e mais três grupos com aplicação de BET), pelo menos uma não foi tomada como amostra da mesma população que as outras distribuições. Em seguida, realizamos os testes de hipótese de Mann Whitney Wilcoxon para amostras independentes e não pareadas, para avaliar a diferença significativa entre os seguintes pares de distribuições: (1) F vs F *; (2) F vs MF; (3) M vs M*; (4) M vs MF; (5) MF vs MF *, onde * significa amostra onde foi aplicada a segmentação BET, M para o sexo masculino, F para o sexo feminino e MF para ambos os sexos biológicos. O objetivo dessas comparações é avaliar se o uso de BET e a separação de gêneros produziram melhores resultados. Os resultados mostram que tanto a aplicação da segmentação BET no pré-processamento quanto a separação de gêneros durante os experimentos trouxeram diferenças significativas aos resultados em todas as comparações acima. Portanto, é necessário fazer comparações específicas entre os resultados.

A Tabela 2 mostra os resultados de precisão global (cenário multiclasse) e os resultados das métricas de precisão, sensibilidade, especificidade, pontuação F1 e precisão para os cenários de classificação binária: (1) indivíduos cognitivamente normais (CN) contra todos; (2) indivíduos com comprometimento cognitivo leve (MCI) contra todos; e indivíduos com Doença de Alzheimer contra todos. Os resultados estão organizados por gênero (feminino, masculino e ambos os gêneros combinados) e foram obtidos sobre conjuntos de imagens que passaram pela segmentação BET como pré-processamento. Considerando a acurácia global, os valores obtidos para mulheres (84,17%), homens (83,59%) e ambos os sexos (77,89%) mostram que sistemas de apoio ao diagnóstico que podem ser construídos considerando a metodologia de escolha das melhores fatias proposta em este trabalho é mais eficiente se eles usam modelos de classificação específicos de gênero. A diminuição no desempenho de precisão quando ambos os sexos são incluídos no mesmo modelo de aprendizado de máquina é de aproximadamente 7%.

Embora se esperasse intuitivamente um resultado de acurácia global para ambos os sexos para aproximar a média aritmética dos resultados para homens e mulheres, isso não ocorre. Como a queda no desempenho é considerável, a hipótese é que diferenças anatômicas nas imagens de ressonância magnética ponderadas em T1 entre cérebros masculinos e femininos de homens devem causar confusão no processo de treinamento da máquina de aprendizado. Isso aponta para a possibilidade de que histórias diferentes para homens e mulheres possam influenciar o processo de envelhecimento cerebral. Para casos de classificação binária, o mesmo vale para cada métrica: os resultados para ambas as classes são de 1% a quase 10% menores que o equivalente para cada gênero separadamente.

Considerando os resultados das classificações binárias de imagens com pré-processamento BET presentes na Tabela 2, temos o seguinte, considerando acurácia, sensibilidade e especificidade. Para o cenário de pacientes sem déficit cognitivo (NC), a acurácia foi de 89,35% (homens) e 90,97% (mulheres), as sensibilidades foram 93,10% (homens) e 94,25% (mulheres), e as especificidades foram 88,79% (homens) e 91,93% (mulheres). Isso pode ser considerado um bom desempenho de discriminação de indivíduos saudáveis (métricas que variam de 90% a 94%, aproximadamente). Para indivíduos com comprometimento cognitivo leve (MCI), a acurácia foi de 87,42% (homens) e 87,74% (mulheres), as sensibilidades foram de 83,16% (para homens e mulheres) e as especificidades foram de 97,50% (homens) e 98,33% (mulheres). Os resultados mostram que, para a discriminação de indivíduos com MCI, há semelhança de desempenho entre os dois modelos de aprendizagem, tanto para homens quanto para mulheres. Os resultados também indicam que, embora os resultados de especificidade sejam muito bons (aproximadamente 98%), o que mostra que os indivíduos que não são classificados como MCI são rejeitados com bastante sucesso, há uma dificuldade um pouco maior em distinguir adequadamente outras classes de MCI (NC e DA). Este é um resultado esperado, uma vez que a classe MCI é uma classe de natureza intermediária entre o indivíduo cognitivamente saudável e aquele com demência mais grave, como a doença de Alzheimer. O valor de Precision para classificação do CCL (ou MCI) foi sempre acima de 98%, o que demonstra que, quando o classificador aponta que é MCI, dificilmente erra. Para os participantes com doença de Alzheimer, DA, a acurácia foi de 93,87% (homens) e 94,52% (mulheres), as sensibilidades foram 90,91% (homens) e 93,92% (mulheres), e as especificidades foram 96,03% (homens) e 96,39% (mulheres). Esses resultados mostram que o modelo de aprendizado de máquina para detecção da doença de Alzheimer foi bastante preciso, com acurácia em torno de 94%, sensibilidades de 91% a 95% e especificidades de 96%. Esses resultados mostram que o mo-

delo de detecção da doença de Alzheimer construído neste trabalho, com pré-processamento com segmentação BET, pode ser considerado preciso. Os resultados foram comparáveis aos trabalhos de última geração, como modelos de aprendizado profundo (BASAIA et al., 2019; LONG et al., 2017; FAROOQ et al., 2017; KOROLEV et al., 2017; MAQSOOD et al., 2019), com 94% para o modelo proposto contra 59-98% para modelos baseados em aprendizado profundo sem seleção de fatias. Isso é muito importante porque, neste trabalho, foram aplicados modelos de aprendizado de máquina de baixo custo computacional sobre atributos relativamente simples baseados em texturas e formas.

A Figura 25 ilustra os histogramas das distribuições das melhores fatias selecionadas pelo algoritmo genético. Esses histogramas são organizados por gênero (mulheres, homens e ambos os sexos) e plano de visão corporal. Os resultados qualitativos mostram que as distribuições de frequência das fatias para homens e mulheres são muito próximas. No entanto, a distribuição das fatias selecionadas considerando ambos os sexos já é um pouco diferente. Nessas três distribuições, entretanto, parece haver duas distribuições normais aparentemente combinadas. A Tabela 4 detalha as distribuições das ocorrências das melhores fatias, mais frequentemente selecionadas, na forma de frequências absolutas e relativas (em porcentagem), organizadas por gênero e por plano corporal. Todos esses resultados foram obtidos a partir da base composta por imagens pré-processadas utilizando o algoritmo de segmentação BET. A Tabela 4 mostra claramente que, para homens e mulheres, as fatias mais frequentemente selecionadas como melhores correspondem ao índice 67-71 para o plano sagital do corpo: 100% das melhores fatias do total de 5030 melhores fatias para mulheres e 100% das melhores fatias do total de 5.041 melhores fatias para homens. Considerando ambos os sexos, o comportamento permanece: os melhores cortes mais frequentemente selecionados são aqueles no plano sagital do corpo, correspondendo aos índices de 66 a 69. Para ambos os sexos, todos os 5.041 melhores cortes mais frequentes foram extraídos da vista sagital. Empiricamente, o córtex entorrinal é melhor visualizado nos planos sagitais (KHAN et al., 2014; HYMAN et al., 1984; GÓMEZ-ISLA et al., 1996; BOBINSKI et al., 1999; ARENDT et al., 2015; MIKKONEN et al., 1999; GIANNAKOPOULOS et al., 2003; ŠIMIĆ et al., 2000; LI et al., 2021). O córtex entorrinal é de fundamental importância para o diagnóstico da doença de Alzheimer. Nossa hipótese é que o córtex entorrinal também foi melhor representado por vetores de atributos extraídos de fatias de planos sagitais. Isso também explicaria a alta precisão obtida para a detecção da doença de Alzheimer.

A Tabela 3 mostra os resultados de precisão global, para a classificação multiclasse (classes CN, MCI e AD), e de especificidade, sensibilidade, precisão, F1-score e precisão, para as

classificações binárias (cada classe contra todos). Os resultados são organizados por gênero: mulheres, homens e ambos os gêneros combinados. Esses resultados correspondem ao cenário de volumes de imagem sem pré-processamento BET. Considerando a acurácia global para o problema de classificação multiclasse, os resultados foram 82,71% (homens), 82,02% (mulheres) e 78,99% (ambos). Observamos que o desempenho é praticamente o mesmo para mulheres e homens, 82%, mas cai um pouco quando as amostras de ambos os sexos são analisadas juntas para cerca de 79%. Essa queda de cerca de 3%, no entanto, não parece estatisticamente significativa. Estes resultados são apenas ligeiramente inferiores aos obtidos no cenário com pré-processamento BET: são cerca de 2% inferiores aos obtidos com BET (cf. Tabela 2), para homens e mulheres separadamente. No entanto, a piora da acurácia para a combinação de ambos os sexos é muito menor no cenário sem BET do que com BET. Considerando indivíduos de cognição normal (CN), as sensibilidades foram 93,42% (homens) e 81,98% (mulheres), as especificidades foram 91,93% (homens) e 94,58% (mulheres), e a acurácia foi de 91,30% (homens) e 86,64% (mulheres). Para ambos os sexos, a sensibilidade foi de 85,56%, a especificidade de 88,69% e a acurácia de 86,46%. Os resultados para ambos os sexos sem BET foram muito próximos daqueles com BET. Enquanto isso, para o cenário BET, notamos que a sensibilidade para homens (cerca de 93%) é razoavelmente maior do que para mulheres (cerca de 82%). Isso permanece para a precisão: cerca de 91% para homens versus cerca de 87% para mulheres. No entanto, no que diz respeito à especificidade, os resultados para os homens, 92%, foram inferiores aos das mulheres, cerca de 95%. Para detectar indivíduos com comprometimento cognitivo leve (MCI), as sensibilidades foram 80,53% (homens) e 89,32% (mulheres), as especificidades foram 98,17% (homens) e 94,83% (mulheres), e a acurácia foi de 86,29% (homens) e 89,89% (mulheres). A acurácia foi praticamente equivalente, de 86% e 90% para homens e mulheres respectivamente, uma diferença de cerca de 4%. Os resultados de sensibilidade são razoavelmente maiores para as mulheres (cerca de 89%) do que para os homens (81%), o que mostra que o reconhecimento de indivíduos com MCI é bastante razoável. Os resultados de especificidade são bastante elevados: cerca de 98% para homens e 95% para mulheres. A diferença de cerca de 3% também mostra que os resultados para homens e mulheres são muito próximos. Os resultados da detecção de MCI para cada gênero separadamente mostram que a rejeição de indivíduos não MCI a serem erroneamente reconhecidos como MCI é alta. Considerando o modelo que combina ambos os sexos, os resultados para sensibilidade, especificidade e acurácia foram 76,11%, 96,47% e 84,90%, respectivamente. Estes resultados são ligeiramente superiores aos obtidos no cenário de pré-processamento com

BET, embora as diferenças de 1% a 2% não pareçam estatisticamente significativas. Para a detecção de indivíduos com doença de Alzheimer (DA), considerando também o cenário sem pré-processamento BET, a acurácia foi de 94,31% (homens) e 90,61% (mulheres), as sensibilidades foram 90,91% (homens) e 92,06 % (mulheres), e as especificidades foram 95,86% (homens) e 92,99% (mulheres). Esses resultados são considerados altos e apenas 1% a 3% inferiores aos obtidos pelos modelos que consideram o cenário com pré-processamento BET (cf. Tabela 2). Para o modelo que combina ambos os sexos, a acurácia, sensibilidade e especificidade foram 89,06%, 87,50% e 90,42%, respectivamente. Esses resultados são muito semelhantes aos do BET, com acurácia, sensibilidade e especificidade de 88,95%, 83,33% e 92,28%, nessa ordem.

A Figura 26 ilustra os histogramas das distribuições das melhores fatias escolhidas pela busca evolutiva, organizadas de acordo com o gênero: feminino (1), masculino (2), ou uma combinação de ambos (3). Esses resultados foram obtidos com imagens sem pré-processamento BET. Cada gráfico possui três histogramas, de acordo com as vistas axial/transversal, coronal ou sagital. Para as mulheres, Figura 26 (1), a maioria absoluta dos melhores cortes corresponde à vista coronal, com alguns sagitais, e aparentemente nenhum axial. A Tabela 5 detalha esses resultados, focando nos melhores cortes mais frequentemente selecionados: 4.490 (97%) do total de 4.629 melhores cortes pertencem à vista coronal; apenas 136 (2,94%) melhores cortes pertencem à vista sagital; e 3 (0,06%) cortes, ou seja, praticamente nenhum corte corresponde ao corte axial. Esses resultados são bem diferentes daqueles obtidos em imagens com BET, onde todos os melhores cortes foram selecionados a partir de visualizações sagitais. Provavelmente as imagens de mulheres na visualização coronal carregam informações anatômicas importantes em regiões que não correspondem a áreas do cérebro. Esta informação deve ter um grande peso nos atributos de textura e forma utilizados neste trabalho para representar as imagens. Dessa forma, isso teria uma grande influência no processo de classificação. Nos histogramas (2) e (3) da Figura 26, podemos ver que o comportamento nas imagens sem pré-processamento BET é mais semelhante ao do cenário com BET: a maioria das melhores fatias selecionadas corresponde a a vista sagital. Na Tabela 5 temos estes resultados mais detalhados. Para os homens, 4662 (97,65%) cortes do total de 4774 melhores cortes correspondem à vista sagital, enquanto 122 (2,35%) cortes são da vista coronal. Esse comportamento é mais semelhante ao cenário com BET. Curiosamente, quando os dois gêneros são combinados para construir um único modelo, os melhores cortes mais frequentes correspondem principalmente à visão sagital: 4613 (98,59%) do total de 4679 cortes. Apenas 66 (1,41%) estavam associados

à visualização coronal. Nenhum dos melhores cortes corresponde à vista axial. Os modelos para o cenário sem BET, ao comparar as métricas presentes nas Tabelas 4 e 5, para a combinação de ambos os gêneros, mostram que os resultados são muito próximos. Os resultados para o cenário sem BET, considerando as métricas de qualidade, mostram que o pré-processamento com BET não torna essas métricas significativamente maiores. Os histogramas das Figuras 25 (3) e 26 (3), e as Tabelas 2 e 3, também mostram que o mais preciso vista é a sagital, e os melhores cortes mais frequentes tendem a coincidir.

6 CONCLUSÃO

6.1 CONCLUSÕES GERAIS

O diagnóstico precoce da doença de Alzheimer é de fundamental importância para melhorar o prognóstico desta doença neurodegenerativa. Buscando contribuir para o desenvolvimento de ferramentas de apoio ao diagnóstico da doença de Alzheimer utilizando imagens de ressonância magnética nuclear estrutural, este trabalho foi motivado pela necessidade de estabelecer uma metodologia para definir os melhores planos de visualização e os melhores cortes de volumes de imagem.

As principais contribuições deste trabalho são: (i) a construção de um método baseado em Computação Evolutiva para selecionar conjuntos de cortes 2D extraídos de um exame de ressonância magnética estrutural que otimize o diagnóstico de doença de Alzheimer e comprometimento cognitivo leve, além da definição de o plano de visualização ideal; (ii) a apresentação de um modelo de sistema de suporte diagnóstico capaz de obter bons resultados utilizando apenas idade, sexo e o conjunto de cortes selecionados automaticamente selecionados a partir do volume de RM do paciente como dados de entrada.

Nossos resultados experimentais mostram que o modelo de aprendizado de máquina para apoiar o diagnóstico da doença de Alzheimer proposto neste trabalho é consideravelmente preciso, atingindo as seguintes métricas de qualidade: acurácia em torno de 94%, sensibilidades de 91-95% e especificidades de 96%. Esses resultados são bastante competitivos quando comparados a vários resultados do estado da arte: 94% para nosso modelo contra 59-98% para modelos baseados em aprendizado profundo sem seleção de fatias. Como empregamos modelos de aprendizado de máquina de baixo custo computacional sobre recursos baseados em texturas e formas, esses resultados são bastante promissores. Além disso, até onde sabemos, não há esforços para automatizar a seleção de fatias em trabalhos de última geração.

Os resultados experimentais foram gerados considerando dois cenários de entrada: com ou sem aplicação de segmentação de extração cerebral como etapa de pré-processamento. Os resultados para ambos os cenários, levando em consideração as métricas de qualidade, demonstram que o pré-processamento de segmentação de extração cerebral não melhora significativamente essas métricas. A visão plana mais precisa é sagital para ambos os cenários, exceto para participantes do sexo feminino sem pré-processamento de segmentação de imagem. As melhores fatias mais frequentes tendem a coincidir. A principal hipótese para sustentar

a indicação do plano sagital como o melhor plano de visualização é a observação empírica de que o córtex entorrinal, estrutura neural fundamental para auxiliar no diagnóstico da doença de Alzheimer, parece ser bem visualizado considerando esse plano de visão.

O modelo proposto pode ser expandido com a inclusão de outros tipos de dados de entrada, como volumes de PET Scan, resultados de questionários cognitivos MSSE e dados genômicos de pacientes relacionados ao gene APOE4. No entanto, nossa proposta baseada em imagens de ressonância magnética ponderadas em T1 pode apoiar o diagnóstico não invasivo de demências com custo razoável e desempenho aceitável, uma vez que a inclusão de PET Scan e dados genômicos poderia aumentar os custos e, conseqüentemente, impedir a adoção dessa técnica na prática clínica.

Por fim, é importante destacar o uso da informação de gênero do paciente. Os modelos especializados em cada gênero trouxeram melhorias significativas nos resultados finais. Este resultado comprova as hipóteses iniciais deste estudo, uma vez que, por diversas razões, a doença de Alzheimer e as demências em geral afetam homens e mulheres de forma diferente, em geral com conseqüências mais graves para as mulheres.

6.2 DIFICULDADES APRESENTADAS

Uma das principais dificuldades enfrentadas foi a falta de detalhes na descrição da metodologia dos trabalhos publicados na área, o que trouxe dificuldades na comparação das soluções propostas e também na identificação de pontos de melhoria. Esse problema na descrição dos detalhes dos trabalhos também dificulta um melhor aproveitamento da grande diversidade de soluções propostas.

Outra dificuldade considerável foi a falta de padronização entre diferentes bases de dados na área, o que dificulta tanto a aderência das soluções propostas como também a interpretação e comparação entre resultados obtidos quando utilizadas diferentes bases.

O início da pandemia de Covid-19 no Brasil em 2020 também impôs duras restrições à realização de diversas atividades humanas. Por conseguinte, ao menos em um primeiro momento, pesquisas científicas que não estavam diretamente ligadas ao combate ao coronavírus precisaram ficar de lado durante um tempo.

Por fim, não poderia ser omitida a dificuldade de conciliação de atividades profissionais na Universidade de Pernambuco, onde atualmente sou servidor público, com as atividades acadêmicas do meu doutorado, nesses últimos anos do percurso do doutorado. Realizar um

doutorado no Brasil sem bolsa de apoio e sem poder dedicar-se integralmente ao projeto ao longo de 4 anos é um desafio considerável.

6.3 CONTRIBUIÇÕES E TRABALHOS FUTUROS

Este trabalho teve como principal contribuição a construção de uma metodologia para selecionar as fatias mais relevantes do ponto de vista da otimização do diagnóstico por imagem em volumes de imagens biomédicas, considerando não somente as fatias obtidas em decorrência do processo de imageamento, mas também fatias devidas à seleção de diferentes planos de visualização, a saber: sagital, coronal e radial. Como estudo de caso, foi escolhido o problema de apoio ao diagnóstico por imagem da doença de Alzheimer e do déficit cognitivo leve, usando volumes de neuroimagens de ressonância magnética. A metodologia proposta se mostrou superior às abordagens baseadas em aprendizado profundo, no que se refere à acurácia, à sensibilidade e à especificidade, que são as métricas mais importantes em aplicações clínicas. O plano de visualização que fornece indicações diagnósticas mais precisas é o sagital, exceto quando se consideram participantes do sexo feminino sem pré-processamento de segmentação de imagem. As melhores fatias mais frequentes tendem a coincidir.

Este trabalho resultou no artigo “Computer-aided diagnosis of Alzheimer’s disease by MRI analysis and evolutionary computing”, publicado na revista *Research on Biomedical Engineering*, Qualis A4 (SOUZA et al., 2021).

Foram também realizadas as seguintes publicações diretamente relacionadas aos resultados deste trabalho:

1. “Deep learning for early diagnosis of Alzheimer’s disease: a contribution and a brief review”, capítulo do livro “Deep Learning for Data Analytics”, publicado pela Elsevier (SILVA et al., 2020a);
2. “Model based on deep feature extraction for diagnosis of Alzheimer’s disease”, publicado na 2019 International Joint Conference on Neural Networks (IJCNN), Qualis A1 (eventos) (SILVA et al., 2019);
3. “Utilização de redes convolucionais para classificação e diagnóstico da doença de Alzheimer”, publicado no II Simpósio de Inovação em Engenharia Biomédica (SILVA et al., 2018).

Adicionalmente, o doutorando participou dos seguintes trabalhos:

1. "Heg.IA: An intelligent system to support diagnosis of Covid-19 based on blood tests", um sistema baseado em aprendizado de máquina para apoio ao diagnóstico rápido e de baixo custo da Covid-19 por meio de resultados de exames de sangue e bioquímicos, publicado na revista *Research on Biomedical Engineering*, Qualis A4 (BARBOSA et al., 2022);
2. "Covid-19 rapid test by combining a Random Forest-based web system and blood tests", versão aperfeiçoada do sistema baseado em aprendizado de máquina para apoio ao diagnóstico rápido e de baixo custo da Covid-19 por meio de resultados de exames de sangue e bioquímicos, publicado no *Journal of Biomolecular Structure and Dynamics*, Qualis A3 (BARBOSA et al., 2021).

Como trabalho futuro, pretendemos investigar o uso de diferentes arquiteturas híbridas baseadas em redes profundas pré-treinadas e classificadores rasos na construção de sistemas inteligentes de apoio ao diagnóstico diferencial para reconhecer imagens de voluntários saudáveis, com a doença de Alzheimer ou com déficit cognitivo leve, a partir das melhores fatias e planos de visualização indicadas neste trabalho.

Adicionalmente, também pretendemos aplicar a metodologia proposta neste trabalho a outros problemas de apoio ao diagnóstico por neuroimagem, considerando imagens PET Scan e de tomografia computadorizada por raios-x, por exemplo. Também se busca investigar a metodologia proposta em problemas de apoio ao diagnóstico de doenças que afetam o trato respiratório, como o câncer de pulmão, a tuberculose e a Covid-19.

REFERÊNCIAS

- ADLER-MILSTEIN, J. From digitization to digital transformation: policy priorities for closing the gap. *JAMA*, American Medical Association, v. 325, n. 8, p. 717–718, 2021.
- ALAM, S.; KWON, G.-R.; INITIATIVE, A. D. N. Alzheimer disease classification using KPCA, LDA, and multi-kernel learning SVM. *International Journal of Imaging Systems and Technology*, Wiley Online Library, v. 27, n. 2, p. 133–143, 2017.
- ALPAYDIN, E. *Introduction to Machine Learning*. 2nd. ed. Cambridge, MA, USA: The MIT Press, 2010. 537 p. ISBN 9780262012430.
- ALPAYDIN, E. *Machine Learning: The New AI (MIT Press Essential Knowledge series)*. The MIT Press, 2016. ISBN 0262529513. Disponível em: <<https://www.amazon.com/Machine-Learning-Press-Essential-Knowledge/dp/0262529513?SubscriptionId=AKIAIOBINVZYXZQZ2U3A&tag=chimbori05-20&linkCode=xm2&camp=2025&creative=165953&creativeASIN=0262529513>>.
- ALTMANN, A.; TIAN, L.; HENDERSON, V. W.; GREICIUS, M. D.; INVESTIGATORS, A. D. N. I. Sex modifies the APOE-related risk of developing Alzheimer's disease. *Annals of Neurology*, Wiley Online Library, v. 75, n. 4, p. 563–573, 2014.
- ARENDET, T.; BRÜCKNER, M. K.; MORAWSKI, M.; JÄGER, C.; GERTZ, H.-J. Early neurone loss in Alzheimer's disease: cortical or subcortical? *Acta Neuropathologica Communications*, Springer, v. 3, n. 1, p. 1–11, 2015.
- ASHFORD, M. T.; NEUHAUS, J.; JIN, C.; CAMACHO, M. R.; FOCKLER, J.; TRURAN, D.; MACKIN, R. S.; RABINOVICI, G. D.; WEINER, M. W.; NOSHENY, R. L. Predicting amyloid status using self-report information from an online research and recruitment registry: The Brain Health Registry. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, Wiley Online Library, v. 12, n. 1, p. e12102, 2020.
- ASIM, Y.; RAZA, B.; MALIK, A. K.; RATHORE, S.; HUSSAIN, L.; IFTIKHAR, M. A. M. A multi-modal, multi-atlas-based approach for Alzheimer detection via machine learning. *International Journal of Imaging Systems and Technology*, v. 28, n. 2, p. 113–123, 2018.
- ASSOCIATION, A. *Beta-amyloid and the amyloid hypothesis*. 2017. Disponível em <http://www.who.int/en/news-room/fact-sheets/detail/dementia>. Acessado em 07/06/2018.
- BALTAXE, E.; CZYPIONKA, T.; KRAUS, M.; REISS, M.; ASKILDSEN, J. E.; GRENKOVIĆ, R.; LINDÉN, T. S.; PITTEr, J. G.; MOLKEN, M. Rutten-van; SOLANS, O. et al. Digital health transformation of integrated care in europe: overarching analysis of 17 integrated care programs. *Journal of Medical Internet Research*, JMIR Publications, v. 21, n. 9, p. e14956, 2019.
- BARBOSA, V. A. d. F.; GOMES, J. C.; SANTANA, M. A. de; LIMA, C. L. de; CALADO, R. B.; JUNIOR, C. R. B.; ALBUQUERQUE, J. E. d. A.; SOUZA, R. G. de; ARAÚJO, R. J. E. de; JUNIOR, L. A. R. M. et al. Covid-19 rapid test by combining a random forest-based web system and blood tests. *Journal of Biomolecular Structure and Dynamics*, Taylor & Francis, p. 1–20, 2021.

BARBOSA, V. A. d. F.; GOMES, J. C.; SANTANA, M. A. de; ALBUQUERQUE, J. E. d. A.; SOUZA, R. G. de; SOUZA, R. E. de; SANTOS, W. P. dos. Heg.IA: An intelligent system to support diagnosis of Covid-19 based on blood tests. *Research on Biomedical Engineering*, Springer, v. 38, n. 1, p. 99–116, 2022.

BASAIA, S.; AGOSTA, F.; WAGNER, L.; CANU, E.; MAGNANI, G.; SANTANGELO, R.; FILIPPI, M. Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. *NeuroImage: Clinical*, Elsevier, v. 21, p. 101645, 2019.

BEN-AHMED, O.; BENOIS-PINEAU, J.; ALLARD, M.; BEN-AMAR, C.; CATHELINE, G. Classification of Alzheimer's disease subjects from MRI using hippocampal visual features. *Multimedia Tools and Applications*, v. 74, n. 4, p. 1249–1266, 2014. ISSN 15737721.

BHATIA, A.; WOLF, E. On the circle polynomials of zernike and related orthogonal sets. In: CAMBRIDGE UNIVERSITY PRESS. *Mathematical Proceedings of the Cambridge Philosophical Society*. [S.l.], 1954. v. 50, n. 1, p. 40–48.

BLESSED, G.; TOMLINSON, B. E.; ROTH, M. The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. *The British Journal of Psychiatry*, Cambridge University Press, v. 114, n. 512, p. 797–811, 1968.

BOBINSKI, M.; LEON, M. J. D.; CONVIT, A.; SANTI, S. D.; WEGIEL, J.; TARSHISH, C. Y.; LOUIS, L. S.; WISNIEWSKI, H. M. MRI of entorhinal cortex in mild Alzheimer's disease. *The Lancet*, Elsevier, v. 353, n. 9146, p. 38–40, 1999.

BOGNER, M. S. *Human error in medicine*. [S.l.]: CRC Press, 2018.

BOSER, B. E.; GUYON, I. M.; VAPNIK, V. N. A training algorithm for optimal margin classifiers. In: ACM. *Proceedings of the fifth annual workshop on Computational learning theory*. [S.l.], 1992. p. 144–152.

BRAVO, G.; HÉBERT, R. Age-and education-specific reference values for the mini-mental and modified mini-mental state examinations derived from a non-demented elderly population. *International Journal of Geriatric Psychiatry*, Wiley Online Library, v. 12, n. 10, p. 1008–1018, 1997.

BRUCKI, S.; NITRINI, R.; CARAMELLI, P.; BERTOLUCCI, P.; OKAMOTO, I. H. Suggestions for utilization of the mini-mental state examination in brazil. *Arquivos de Neuro-Psiquiatria*, v. 61, n. 3B, p. 777–781, 2003.

BUNKHUMPORNPAT, C.; SINAPIROMSARAN, K.; LURSINSAP, C. Safe-level-smote: Safe-level-synthetic minority over-sampling technique for handling the class imbalanced problem. In: SPRINGER. *Pacific-Asia Conference on Knowledge Discovery and Data Mining*. [S.l.], 2009. p. 475–482.

BUTTERS, N.; GRANHOLM, E.; SALMON, D. P.; GRANT, I.; WOLFE, J. Episodic and semantic memory: A comparison of amnesic and demented patients. *Journal of Clinical and Experimental Neuropsychology*, Taylor & Francis, v. 9, n. 5, p. 479–497, 1987.

CAHN, D. A.; SALMON, D. P.; MONSCH, A. U.; BUTTERS, N.; WIEDERHOLT, W.; COREY-BLOOM, J.; BARRETT-CONNOR, E. Screening for dementia of the alzheimer type in the community: the utility of the clock drawing test. *Archives of Clinical Neuropsychology*, Elsevier, v. 11, n. 6, p. 529–539, 1996.

CARDOSO, B. R.; HARE, D. J.; BUSH, A. I.; LI, Q.-X.; FOWLER, C. J.; MASTERS, C. L.; MARTINS, R. N.; GANIO, K.; LOTHIAN, A.; MUKHERJEE, S. et al. Selenium levels in serum, red blood cells, and cerebrospinal fluid of Alzheimer's disease patients: a report from the Australian Imaging, Biomarker & Lifestyle Flagship Study of Ageing (AIBL). *Journal of Alzheimer's Disease*, IOS press, v. 57, n. 1, p. 183–193, 2017.

CASELLI, R. J.; BEACH, T. G.; YAARI, R.; REIMAN, E. M. Alzheimer's disease a century later. *The Journal of Clinical Psychiatry*, Physicians Postgraduate Press, 2006.

CHAWLA, N. V.; BOWYER, K. W.; HALL, L. O.; KEGELMEYER, W. P. SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, v. 16, p. 321–357, 2002.

CHEN, Y. J.; DEUTSCH, G.; SATYA, R.; LIU, H.-G.; MOUNTZ, J. M. A semi-quantitative method for correlating brain disease groups with normal controls using spect: Alzheimer's disease versus vascular dementia. *Computerized Medical Imaging and Graphics*, Elsevier, v. 37, n. 1, p. 40–47, 2013.

CHUNG, B. S.; PARK, J. S. Automatic segmentation of true color sectioned images using fmrib software library: First trial in brain, gray matter, and white matter. *Clinical Anatomy*, Wiley Online Library, v. 33, n. 8, p. 1197–1203, 2020.

CHUTE, C.; FRENCH, T. Introducing care 4.0: an integrated care paradigm built on industry 4.0 capabilities. *International journal of environmental research and public health*, Multidisciplinary Digital Publishing Institute, v. 16, n. 12, p. 2247, 2019.

COELHO, L. P. Mahotas: Open source software for scriptable computer vision. *Journal of Open Research Software*, v. 1, 2013. Disponível em: <<http://dx.doi.org/10.5334/jors.ac>>.

COMMOWICK, O.; ISTACE, A.; KAIN, M.; LAURENT, B.; LERAY, F.; SIMON, M.; POP, S. C.; GIRARD, P.; AMELI, R.; FERRÉ, J.-C. et al. Objective evaluation of multiple sclerosis lesion segmentation using a data management and processing infrastructure. *Scientific Reports*, Nature Publishing Group, v. 8, n. 1, p. 1–17, 2018.

CORDEIRO, F. R.; SANTOS, W. P.; SILVA-FILHO, A. G. A semi-supervised fuzzy growcut algorithm to segment and classify regions of interest of mammographic images. *Expert Systems with Applications*, Elsevier, v. 65, p. 116–126, 2016.

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

CUMMINGS, J. L. The neuropsychiatric inventory: assessing psychopathology in dementia patients. *Neurology*, AAN Enterprises, v. 48, n. 5 Suppl 6, p. 10S–16S, 1997.

CUMMINGS, J. L.; MEGA, M.; GRAY, K.; ROSENBERG-THOMPSON, S.; CARUSI, D. A.; GORNBEIN, J. The neuropsychiatric inventory: comprehensive assessment of psychopathology in dementia. *Neurology*, AAN Enterprises, v. 44, n. 12, p. 2308–2308, 1994.

DICE, L. R. Measures of the amount of ecologic association between species. *Ecology*, Wiley Online Library, v. 26, n. 3, p. 297–302, 1945.

DICKIE, D. A.; SHENKIN, S. D.; ANBLAGAN, D.; LEE, J.; CABEZ, M. B.; RODRIGUEZ, D.; BOARDMAN, J. P.; WALDMAN, A.; JOB, D. E.; WARDLAW, J. M. Whole brain magnetic resonance image atlases: a systematic review of existing atlases and caveats for use in population imaging. *Frontiers in Neuroinformatics*, Frontiers, v. 11, p. 1, 2017.

DIETTERICH, T. G. *Machine learning in ecosystem informatics and sustainability*. [S.l.: s.n.], 2009. 8–13 p. ISSN 10450823. ISBN 9781577354260.

DOUZAS, G.; BACAO, F. Geometric SMOTE a geometrically enhanced drop-in replacement for SMOTE. *Information Sciences*, Elsevier, v. 501, p. 118–135, 2019.

DOUZAS, G.; BACAO, F.; LAST, F. Improving imbalanced learning through a heuristic oversampling method based on k-means and SMOTE. *Information Sciences*, Elsevier, v. 465, p. 1–20, 2018.

DRUCKER, H.; WU, D.; VAPNIK, V. N. Support vector machines for spam categorization. *IEEE Transactions on Neural networks*, IEEE, v. 10, n. 5, p. 1048–1054, 1999.

DZAU, V. J.; BALATBAT, C. A. Future of hypertension: The need for transformation. *Hypertension*, Am Heart Assoc, v. 74, n. 3, p. 450–457, 2019.

EIKELBOOM, W. S.; PAN, M.; OSSENKOPPELE, R.; COESMANS, M.; GATCHEL, J. R.; ISMAIL, Z.; LANCTÔT, K. L.; FISCHER, C. E.; MORTBY, M. E.; BERG, E. van den et al. Sex differences in neuropsychiatric symptoms in Alzheimer's disease dementia: a meta-analysis. *Alzheimer's Research & Therapy*, BioMed Central, v. 14, n. 1, p. 1–13, 2022.

ELBEJJANI, M.; FUHRER, R.; ABRAHAMOWICZ, M.; MAZOYER, B.; CRIVELLO, F.; TZOURIO, C.; DUFOUIL, C. Depression, depressive symptoms, and rate of hippocampal atrophy in a longitudinal cohort of older men and women. *Psychological Medicine*, Cambridge University Press, v. 45, n. 9, p. 1931, 2015.

ELLIS, K. A.; BUSH, A. I.; DARBY, D.; FAZIO, D. D.; FOSTER, J.; HUDSON, P.; LAUTENSCHLAGER, N. T.; LENZO, N.; MARTINS, R. N.; MARUFF, P. et al. The Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging: methodology and baseline characteristics of 1112 individuals recruited for a longitudinal study of alzheimer's disease. *International Psychogeriatrics*, Cambridge University Press, v. 21, n. 4, p. 672–687, 2009.

ELLIS, K. A.; SZOEKE, C.; BUSH, A. I.; DARBY, D.; GRAHAM, P. L.; LAUTENSCHLAGER, N. T.; MACAULAY, S. L.; MARTINS, R. N.; MARUFF, P.; MASTERS, C. L. et al. Rates of diagnostic transition and cognitive change at 18-month follow-up among 1,112 participants in the Australian Imaging, Biomarkers and Lifestyle Flagship Study of Ageing (AIBL). *International Psychogeriatrics*, Cambridge University Press, v. 26, n. 4, p. 543–554, 2014.

EZHILARASAN, K.; PRAVEENKUMAR, S.; SOMASUNDARAM, K.; KALAISELVI, T.; MAGESH, S.; KIRUTHIKA, S.; JEEVAREKHA, A. Automatic brain extraction from mri of human head scans using helmholtz free energy principle and morphological operations. *Biomedical Signal Processing and Control*, Elsevier, v. 64, p. 102270, 2021.

- FANG, C.; LI, C.; CABRERIZO, M.; BARRETO, A.; ANDRIAN, J.; LOEWENSTEIN, D.; DUARA, R.; ADJOUADI, M. A novel Gaussian discriminant analysis-based computer aided diagnosis system for screening different stages of Alzheimer's disease. In: IEEE. *2017 IEEE 17th International Conference on Bioinformatics and Bioengineering (BIBE)*. [S.l.], 2017. p. 279–284.
- FARIAS, S. T.; MUNGAS, D.; REED, B. R.; CAHN-WEINER, D.; JAGUST, W.; BAYNES, K.; DECARLI, C. The measurement of everyday cognition (ECog): scale development and psychometric properties. *Neuropsychology*, American Psychological Association, v. 22, n. 4, p. 531, 2008.
- FAROOQ, A.; ANWAR, S.; AWAIS, M.; REHMAN, S. A deep CNN based multi-class classification of Alzheimer's disease using mri. In: IEEE. *2017 IEEE International Conference on Imaging systems and techniques (IST)*. [S.l.], 2017. p. 1–6.
- FARRER, L. A.; CUPPLES, L. A.; HAINES, J. L.; HYMAN, B.; KUKULL, W. A.; MAYEUX, R.; MYERS, R. H.; PERICAK-VANCE, M. A.; RISCH, N.; DUIJN, C. M. V. Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer's disease: a meta-analysis. *JAMA*, American Medical Association, v. 278, n. 16, p. 1349–1356, 1997.
- FERNÁNDEZ, A.; GARCIA, S.; HERRERA, F.; CHAWLA, N. V. SMOTE for learning from imbalanced data: progress and challenges, marking the 15-year anniversary. *Journal of Artificial Intelligence Research*, v. 61, p. 863–905, 2018.
- FOLSTEIN, M. F.; ROBINS, L. N.; HELZER, J. E. The mini-mental state examination. *Archives of General Psychiatry*, American Medical Association, v. 40, n. 7, p. 812–812, 1983.
- FREIRE, A. M. A. Pedagogia da indignação: cartas pedagógicas e outros escritos. *Interface-Comunicação, Saúde, Educação*, SciELO Brasil, v. 5, p. 147–152, 2001.
- FREITAS, S.; SIMÕES, M. R.; ALVES, L.; SANTANA, I. Montreal cognitive assessment: validation study for mild cognitive impairment and alzheimer disease. *Alzheimer Disease & Associated Disorders*, LWW, v. 27, n. 1, p. 37–43, 2013.
- GALWAY, L.; CHARLES, D.; BLACK, M. Machine learning in digital games: a survey. *Artificial Intelligence Review*, v. 29, n. 2, p. 123–161, apr 2008. ISSN 0269-2821. Disponível em: <<http://link.springer.com/10.1007/s10462-009-9112-y>>.
- GIANNAKOPOULOS, P.; HERRMANN, F.; BUSSIÈRE, T.; BOURAS, C.; KÖVARI, E.; PERL, D.; MORRISON, J.; GOLD, G.; HOF, P. Tangle and neuron numbers, but not amyloid load, predict cognitive status in Alzheimer's disease. *Neurology*, AAN Enterprises, v. 60, n. 9, p. 1495–1500, 2003.
- GOLLAN, T. H.; WEISSBERGER, G. H.; RUNNQVIST, E.; MONTOYA, R. I.; CERA, C. M. Self-ratings of spoken language dominance: A multi-lingual naming test (mint) and preliminary norms for young and aging spanish-english bilinguals. *Bilingualism*, NIH Public Access, Cambridge, England, v. 15, n. 3, p. 594, 2012.
- GOMES, J. C.; BARBOSA, V. A.; RIBEIRO, D. E.; SOUZA, R. E. de; SANTOS, W. P. dos. Electrical impedance tomography image reconstruction based on backprojection and extreme learning machines. *Research on Biomedical Engineering*, Springer, v. 2020, n. 36, p. 399—410, 2020.

GOMES, J. C.; BARBOSA, V. A. de F.; SANTANA, M. A. de; BANDEIRA, J.; VALENCA, M. J. S.; SOUZA, R. E. de; ISMAEL, A. M.; SANTOS, W. P. dos. Ikonos: An intelligent tool to support diagnosis of covid-19 by texture analysis of x-ray images. *medRxiv*, Cold Spring Harbor Laboratory Press, 2020.

GOMES, J. C.; PEREIRA, J. M. S.; SANTANA, M. A. de; SILVA, W. W. A. da; SOUZA, R. E. de; SANTOS, W. P. dos. Electrical impedance tomography image reconstruction based on autoencoders and extreme learning machines. In: *Deep Learning for Data Analytics: Foundations, Biomedical Applications, and Challenges*. [S.l.]: Elsevier, 2020. p. 155–171.

GÓMEZ-ISLA, T.; PRICE, J. L.; JR, D. W. M.; MORRIS, J. C.; GROWDON, J. H.; HYMAN, B. T. Profound loss of layer II entorhinal cortex neurons occurs in very mild Alzheimer's disease. *Journal of Neuroscience*, Soc Neuroscience, v. 16, n. 14, p. 4491–4500, 1996.

GONZALEZ, R. C.; WOODS, R. E. *Digital image processing*. Upper Saddle River, N.J.: Prentice Hall, 2008. ISBN 9780131687288 013168728X 9780135052679 013505267X. Disponível em: <<http://www.amazon.com/Digital-Image-Processing-3rd-Edition/dp/013168728X>>.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep Learning*. [S.l.]: MIT Press, 2016. <<http://www.deeplearningbook.org>>.

GORJI, H. T.; HADDADNIA, J. A novel method for early diagnosis of Alzheimer's disease based on pseudo Zernike moment from structural MRI. *Neuroscience*, v. 305, p. 361–371, 2015. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-84941284847&doi=10.1016/j.neuroscience.2015.08.013&partnerID=40&md5=f1d860639b473d8b0d0a>>.

GRAY, K. R.; WOLZ, R.; HECKEMANN, R. A.; ALJABAR, P.; HAMMERS, A.; RUECKERT, D.; INITIATIVE, A. D. N. et al. Multi-region analysis of longitudinal fdg-pet for the classification of alzheimer's disease. *NeuroImage*, Elsevier, v. 60, n. 1, p. 221–229, 2012.

GROBER, E.; SLIWINSKI, M.; KOREY, S. R. Development and validation of a model for estimating premorbid verbal intelligence in the elderly. *Journal of Clinical and Experimental Neuropsychology*, Taylor & Francis, v. 13, n. 6, p. 933–949, 1991.

GUNTER, J.; THOSTENSON, K.; BOROWSKI, B.; REID, R.; ARANI, A.; BERNSTEIN, M.; FOX, N.; THOMAS, D.; DECARLI, C.; TOSUN, D. et al. ADNI-3 MRI protocol. *Alzheimer's & Dementia*, v. 13, n. 7, p. P104–P105, 2017.

HAMDAN, A. C. Avaliação neuropsicológica na doença de Alzheimer e no comprometimento cognitivo leve. *Psicologia Argumento*, v. 26, n. 54, p. 183–192, 2008.

HAN, H.; WANG, W.-Y.; MAO, B.-H. Borderline-SMOTE: a new over-sampling method in imbalanced data sets learning. In: SPRINGER. *International Conference on Intelligent Computing*. [S.l.], 2005. p. 878–887.

HANYU, H.; SATO, T.; HIRAO, K.; KANETAKA, H.; IWAMOTO, T.; KOIZUMI, K. The progression of cognitive deterioration and regional cerebral blood flow patterns in alzheimer's disease: a longitudinal spect study. *Journal of the neurological sciences*, Elsevier, v. 290, n. 1-2, p. 96–101, 2010.

HARALICK, R. M.; DINSTEIN, I.; SHANMUGAM, K. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, IEEE, v. 3, n. 6, p. 610–621, 1973. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4309314>>.

HARDY, J. Alzheimer's disease: the amyloid cascade hypothesis: an update and reappraisal. *Journal of Alzheimer's disease*, IOS Press, v. 9, n. s3, p. 151–153, 2006.

HAYKIN, S. O. *Neural Networks and Learning Machines (3rd Edition)*. Pearson, 2008. ISBN 0131471392. Disponível em: <<https://www.amazon.com/Neural-Networks-Learning-Machines-3rd/dp/0131471392?SubscriptionId=AKIAIOBINVZYXZQZ2U3A&tag=chimbori05-20&linkCode=xm2&camp=2025&creative=165953&creative>>.

HECKEMANN, R. A.; LEDIG, C.; GRAY, K. R.; ALJABAR, P.; RUECKERT, D.; HAJNAL, J. V.; HAMMERS, A. Brain extraction using label propagation and group agreement: Pincram. *PLoS One*, Public Library of Science, v. 10, n. 7, p. e0129211, 2015.

HENDERSON, V. W.; BUCKWALTER, J. G. Cognitive deficits of men and women with Alzheimer's disease. *Neurology*, AAN Enterprises, v. 44, n. 1, p. 90–90, 1994.

HODGES, J. R. Alzheimer's centennial legacy: origins, landmarks and the current status of knowledge concerning cognitive aspects. *Brain*, Oxford University Press, v. 129, n. 11, p. 2811–2822, 2006.

HÖRDER, H.; JOHANSSON, L.; GUO, X.; GRIMBY, G.; KERN, S.; ÖSTLING, S.; SKOOG, I. Midlife cardiovascular fitness and dementia: a 44-year longitudinal population study in women. *Neurology*, AAN Enterprises, v. 90, n. 15, p. e1298–e1305, 2018.

HUANG, H.-C.; TSENG, Y.-M.; CHEN, Y.-C.; CHEN, P.-Y.; CHIU, H.-Y. Diagnostic accuracy of the clinical dementia rating scale for detecting mild cognitive impairment and dementia: A bivariate meta-analysis. *International Journal of Geriatric Psychiatry*, Wiley Online Library, v. 36, n. 2, p. 239–251, 2021.

HUGHES, C. P.; BERG, L.; DANZIGER, W.; COBEN, L. A.; MARTIN, R. L. A new clinical scale for the staging of dementia. *The British Journal of Psychiatry*, Cambridge University Press, v. 140, n. 6, p. 566–572, 1982.

HYMAN, B. T.; HOESEN, G. W. V.; DAMASIO, A. R.; BARNES, C. L. Alzheimer's disease: cell-specific pathology isolates the hippocampal formation. *Science*, American Association for the Advancement of Science, v. 225, n. 4667, p. 1168–1170, 1984.

ISENSEE, F.; SCHELL, M.; PFLUEGER, I.; BRUGNARA, G.; BONEKAMP, D.; NEUBERGER, U.; WICK, A.; SCHLEMMER, H.-P.; HEILAND, S.; WICK, W. et al. Automated brain extraction of multisequence MRI using artificial neural networks. *Human Brain Mapping*, Wiley Online Library, v. 40, n. 17, p. 4952–4964, 2019.

ISLAM, J.; ZHANG, Y. An ensemble of deep convolutional neural networks for Alzheimer's disease detection and classification. *arXiv preprint arXiv:1712.01675*, 2017.

JACK-JR, C. R.; KNOPMAN, D. S.; JAGUST, W. J.; SHAW, L. M.; AISEN, P. S.; WEINER, M. W.; PETERSEN, R. C.; TROJANOWSKI, J. Q. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *The Lancet Neurology*, Elsevier, v. 9, n. 1, p. 119–128, 2010.

- JAYARAMAN, P. P.; FORKAN, A. R. M.; MORSHED, A.; HAGHIGHI, P. D.; KANG, Y.-B. Healthcare 4.0: A review of frontiers in digital health. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, Wiley Online Library, v. 10, n. 2, p. e1350, 2020.
- JENKINSON, M.; PECHAUD, M.; SMITH, S. et al. BET2: MR-based estimation of brain, skull and scalp surfaces. In: *Eleventh annual meeting of the organization for human brain mapping*. [S.l.: s.n.], 2005. v. 17, p. 167.
- JR, C. R. J.; BERNSTEIN, M. A.; FOX, N. C.; THOMPSON, P.; ALEXANDER, G.; HARVEY, D.; BOROWSKI, B.; BRITSON, P. J.; WHITWELL, J. L.; WARD, C. 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.
- KAUFER, D. I.; CUMMINGS, J. L.; KETCHEL, P.; SMITH, V.; MACMILLAN, A.; SHELLEY, T.; LOPEZ, O. L.; DEKOSKY, S. T. Validation of the NPI-Q, a brief clinical form of the Neuropsychiatric Inventory. *The Journal of Neuropsychiatry and Clinical Neurosciences*, Am Neuropsych Assoc, v. 12, n. 2, p. 233–239, 2000.
- KHAN, U. A.; LIU, L.; PROVENZANO, F. A.; BERMAN, D. E.; PROFACI, C. P.; SLOAN, R.; MAYEUX, R.; DUFF, K. E.; SMALL, S. A. Molecular drivers and cortical spread of lateral entorhinal cortex dysfunction in preclinical Alzheimer's disease. *Nature Neuroscience*, Nature Publishing Group, v. 17, n. 2, p. 304–311, 2014.
- KHOTANZAD, A.; HONG, Y. H. Invariant image recognition by Zernike moments. *IEEE Transactions on pattern analysis and machine intelligence*, IEEE, v. 12, n. 5, p. 489–497, 1990.
- KIM, J. W.; BYUN, M. S.; SOHN, B. K.; YI, D.; SEO, E. H.; CHOE, Y. M.; KIM, S. G.; CHOI, H. J.; LEE, J. H.; CHEE, I. S.; WOO, J. I.; LEE, D. Y. Clinical dementia rating orientation score as an excellent predictor of the progression to Alzheimer's disease in mild cognitive impairment. *Psychiatry investigation*, Korean Neuropsychiatric Association, v. 14, n. 4, p. 420, 2017.
- KOROLEV, S.; SAFIULLIN, A.; BELYAEV, M.; DODONOVA, Y. Residual and plain convolutional neural networks for 3D brain MRI classification. In: IEEE. *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*. [S.l.], 2017. p. 835–838.
- LAMA, R. K.; GWAK, J.; PARK, J.-S.; LEE, S.-W. Diagnosis of alzheimer's disease based on structural mri images using a regularized extreme learning machine and pca features. *Journal of healthcare engineering*, Hindawi, v. 2017, 2017.
- LAWS, K. R.; IRVINE, K.; GALE, T. M. Sex differences in cognitive impairment in Alzheimer's disease. *World Journal of Psychiatry*, Baishideng Publishing Group Inc, v. 6, n. 1, p. 54, 2016.
- LEE, J. G.; HAN, J.; HONG, M. H.; WON, D. J.; LEE, S. Y.; LEE, J. Y.; KI, E. C.; KWEON, D. C. Analysis of human error on medical imaging in mass media. *Journal of Radiation Industry*, v. 11, n. 4, p. 259–263, 2018.
- LEUNG, K. K.; BARNES, J.; MODAT, M.; RIDGWAY, G. R.; BARTLETT, J. W.; FOX, N. C.; OURSELIN, S.; INITIATIVE, A. D. N. et al. Brain MAPS: an automated, accurate

and robust brain extraction technique using a template library. *NeuroImage*, Elsevier, v. 55, n. 3, p. 1091–1108, 2011.

LI, F.; CHENG, D.; LIU, M. Alzheimer's disease classification based on combination of multi-model convolutional networks. In: IEEE. *2017 IEEE International Conference on Imaging Systems and Techniques (IST)*. [S.l.], 2017. p. 1–5.

LI, Q.; ZHANG, Y.; LIANG, H.; GONG, H.; JIANG, L.; LIU, Q.; SHEN, L. Deep learning based neuronal soma detection and counting for alzheimer's disease analysis. *Computer Methods and Programs in Biomedicine*, Elsevier, v. 203, p. 106023, 2021.

LIMA, S. M. de; SILVA-FILHO, A. G. da; SANTOS, W. P. dos. Detection and classification of masses in mammographic images in a multi-kernel approach. *Computer methods and programs in biomedicine*, Elsevier, v. 134, p. 11–29, 2016.

LIMA, S. M. L. de; SILVA-FILHO, A. G. da; SANTOS, W. P. dos. Detection and classification of masses in mammographic images in a multi-kernel approach. *Computer Methods and Programs in Biomedicine*, v. 134, p. 11–29, 2016. ISSN 0169-2607. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0169260716304242>>.

LIU, J.; LI, M.; LAN, W.; WU, F.-X.; PAN, Y.; WANG, J. 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.

LONG, X.; CHEN, L.; JIANG, C.; ZHANG, L.; INITIATIVE, A. D. N. Prediction and classification of Alzheimer disease based on quantification of MRI deformation. *PLoS One*, Public Library of Science San Francisco, CA USA, v. 12, n. 3, p. e0173372, 2017.

LORENA, A.; CARVALHO, A. Uma introdução às support vector machines. v. 14, 01 2007.

LOW, W. C. S.; CHUAH, J. H.; TEE, C. A. T.; ANIS, S.; SHOAIB, M. A.; FAISAL, A.; KHALIL, A.; LAI, K. W. An overview of deep learning techniques on chest x-ray and ct scan identification of covid-19. *Computational and Mathematical Methods in Medicine*, Hindawi, v. 2021, 2021.

LUCAS, S. M. Computational intelligence and games: Challenges and opportunities. *International Journal of Automation and Computing*, v. 5, n. 1, p. 45–57, jan 2008. ISSN 1476-8186. Disponível em: <<http://link.springer.com/10.1007/s11633-008-0045-8>>.

MACIEJEWSKI, T.; STEFANOWSKI, J. Local neighbourhood extension of SMOTE for mining imbalanced data. In: IEEE. *2011 IEEE Symposium on Computational Intelligence and Data Mining (CIDM)*. [S.l.], 2011. p. 104–111.

MALLOY-DINIZ, L. F.; LASMAR, V. A. P.; GAZINELLI, L. d. S. R.; FUENTES, D.; SALGADO, J. V. The rey auditory-verbal learning test: applicability for the brazilian elderly population. *Brazilian Journal of Psychiatry*, SciELO Brasil, v. 29, n. 4, p. 324–329, 2007.

MANNING, C. D.; RAGHAVAN, P.; SCHÜTZE, H. *Introduction to Information Retrieval*. Cambridge University Press, 2008. ISBN 0521865719. Disponível em: <<https://www.amazon.com/Introduction-Information-Retrieval-Christopher-Manning/dp/0521865719?SubscriptionId=AKIAIOBINVZYXZQZ2U3A&tag=chimbori05-20&linkCode=xm2&camp=2025&creative=165953&creativeASIN=0521865719>>.

MANUAL, W. D. W. M. S. *revised*. San Antonio, USA: Psychological Corporation Harcourt Brace Jovanovich, 1987.

MAQSOOD, M.; NAZIR, F.; KHAN, U.; AADIL, F.; JAMAL, H.; MEHMOOD, I.; SONG, O.-y. Transfer learning assisted classification and detection of Alzheimer's disease stages using 3D MRI scans. *Sensors*, Multidisciplinary Digital Publishing Institute, v. 19, n. 11, p. 2645, 2019.

MARSHALL, G. A.; OLSON, L. E.; FREY, M. T.; MAYE, J.; BECKER, J. A.; RENTZ, D. M.; SPERLING, R. A.; JOHNSON, K. A.; INITIATIVE, A. D. N. et al. Instrumental activities of daily living impairment is associated with increased amyloid burden. *Dementia and geriatric cognitive disorders*, Karger Publishers, v. 31, n. 6, p. 443–450, 2011.

MARSON, D. Investigating functional impairment in preclinical Alzheimer's disease. *The Journal of Prevention of Alzheimer's Disease*, NIH Public Access, v. 2, n. 1, p. 4, 2015.

MARTINKOVA, J.; QUEVENCO, F.-C.; KARCHER, H.; FERRARI, A.; SANDSET, E. C.; SZOEKE, C.; HORT, J.; SCHMIDT, R.; CHADHA, A. S.; FERRETTI, M. T. Proportion of women and reporting of outcomes by sex in clinical trials for Alzheimer Disease: a systematic review and meta-analysis. *JAMA Network Open*, American Medical Association, v. 4, n. 9, p. e2124124–e2124124, 2021.

MARUFF, P.; LIM, Y. Y.; DARBY, D.; ELLIS, K. A.; PIETRZAK, R. H.; SNYDER, P. J.; BUSH, A. I.; SZOEKE, C.; SCHEMBRI, A.; AMES, D. et al. Clinical utility of the cogstate brief battery in identifying cognitive impairment in mild cognitive impairment and alzheimer's disease. *BMC Psychology*, BioMed Central, v. 1, n. 1, p. 1–11, 2013.

MARUFF, P.; THOMAS, E.; CYSIQUE, L.; BREW, B.; COLLIE, A.; SNYDER, P.; PIETRZAK, R. H. Validity of the cogstate brief battery: relationship to standardized tests and sensitivity to cognitive impairment in mild traumatic brain injury, schizophrenia, and aids dementia complex. *Archives of Clinical Neuropsychology*, Oxford University Press, v. 24, n. 2, p. 165–178, 2009.

MATTHEWS, F. E.; STEPHAN, B. C.; MCKEITH, I. G.; BOND, J.; BRAYNE, C.; FUNCTION, M. R. C. C.; STUDY, A. Two-year progression from mild cognitive impairment to dementia: to what extent do different definitions agree? *Journal of the American Geriatrics Society*, Wiley Online Library, v. 56, n. 8, p. 1424–1433, 2008.

MIELKE, M. M.; VEMURI, P.; ROCCA, W. A. Clinical epidemiology of Alzheimer's disease: assessing sex and gender differences. *Clinical epidemiology*, Dove Press, v. 6, p. 37, 2014.

MIKKONEN, M.; ALAFUZOFF, I.; TAPIOLA, T.; SOININEN, H.; MIETTINEN, R. Subfield- and layer-specific changes in parvalbumin, calretinin and calbindin-D28K immunoreactivity in the entorhinal cortex in Alzheimer's disease. *Neuroscience*, Elsevier, v. 92, n. 2, p. 515–532, 1999.

MITCHELL, A. J. The mini-mental state examination (mmse): update on its diagnostic accuracy and clinical utility for cognitive disorders. In: *Cognitive screening instruments*. [S.l.]: Springer, 2017. p. 37–48.

MITCHELL, T. M. *Machine Learning*. McGraw-Hill Education, 1997. ISBN 0070428077. Disponível em: <<https://www.amazon.com/Machine-Learning-Tom-M-Mitchell/dp/>>

0070428077?SubscriptionId=AKIAIOBINVZYXZQZ2U3A&tag=chimbori05-20&linkCode=xm2&camp=2025&creative=165953&creativeASIN=0070428077>.

MORRIS, J. Current vision and scoring rules the clinical dementia rating (CDR). *Neurology*, v. 43, p. 2412–2414, 1993.

MUELLER, S. G.; WEINER, M. W.; THAL, L. J.; PETERSEN, R. C.; JACK, C. R.; JAGUST, W.; TROJANOWSKI, J. Q.; TOGA, A. W.; BECKETT, L. Ways toward an early diagnosis in Alzheimer's disease: the Alzheimer's disease neuroimaging initiative (ADNI). *Alzheimer's & Dementia*, Elsevier, v. 1, n. 1, p. 55–66, 2005.

NASREDDINE, Z.; COLLIN, I.; CHERTKOW, H.; PHILLIPS, N.; BERGMAN, H.; WHITEHEAD, V. Sensitivity and specificity of the montreal cognitive assessment (moca) for detection of mild cognitive deficits. *Can J Neurol Sci*, v. 30, n. 2, p. 30, 2003.

NELSON, H. E.; O'CONNELL, A. Dementia: the estimation of premorbid intelligence levels using the new adult reading test. *Cortex*, Elsevier, v. 14, n. 2, p. 234–244, 1978.

OKONKWO, O. C.; ALOSCO, M. L.; GRIFFITH, H. R.; MIELKE, M. M.; SHAW, L. M.; TROJANOWSKI, J. Q.; TREMONT, G.; INITIATIVE, A. D. N. et al. Cerebrospinal fluid abnormalities and rate of decline in everyday function across the dementia spectrum: Normal aging, mild cognitive impairment, and Alzheimer disease. *Archives of Neurology*, v. 67, n. 6, p. 688–696, 2010. Disponível em: <+http://dx.doi.org/10.1001/archneurol.2010.118>.

OTT, B. R.; TATE, C. A.; GORDON, N. M.; HEINDEL, W. C. Gender differences in the behavioral manifestations of alzheimer's disease. *Journal of the American Geriatrics Society*, Wiley Online Library, v. 44, n. 5, p. 583–587, 1996.

PETERSEN, R. C. Mild cognitive impairment as a diagnostic entity. *Journal of internal medicine*, Wiley Online Library, v. 256, n. 3, p. 183–194, 2004.

PETERSEN, R. C.; AISEN, P.; BECKETT, L. A.; DONOHUE, M.; GAMST, A.; HARVEY, D. J.; JACK, C.; JAGUST, W.; SHAW, L.; TOGA, A. et al. Alzheimer's disease neuroimaging initiative (ADNI): clinical characterization. *Neurology*, AAN Enterprises, v. 74, n. 3, p. 201–209, 2010.

PETERSEN, R. C.; SMITH, G. E.; WARING, S. C.; IVNIK, R. J.; TANGALOS, E. G.; KOKMEN, E. Mild cognitive impairment: clinical characterization and outcome. *Archives of neurology*, American Medical Association, v. 56, n. 3, p. 303–308, 1999.

PETRACCA, F.; CIANI, O.; CUCCINIELLO, M.; TARRICONE, R. et al. Harnessing digital health technologies during and after the COVID-19 pandemic: context matters. *Journal of medical Internet research*, JMIR Publications, v. 22, n. 12, p. e21815, 2020.

PFEFFER, R. I.; KUROSAKI, T. T.; JR, C. H.; CHANCE, J. M.; FILOS, S. Measurement of functional activities in older adults in the community. *Journal of Gerontology*, The Gerontological Society of America, v. 37, n. 3, p. 323–329, 1982.

POPESCU, V.; M.BATTAGLINI; HOOGSTRATE, W. S.; VERFAILLIE, S.; SLUIMER, I. C.; SCHIJNDEL, R. A. van; DIJK, B. W. van; COVER, K. S.; KNOL, D. L.; JENKINSON, M.; BARKHOF, F.; STEFANO, N. de; VRENKEN, H. Optimizing parameter choice for FSL-Brain Extraction Tool (BET) on 3D T1 images in multiple sclerosis. *NeuroImage*, Elsevier, v. 61, n. 4, p. 1484–1494, 2012.

PROMTEANGTRONG, C.; KOLBER, M.; RAMCHANDRA, P.; MOGHBEL, M.; HOUSHMAND, S.; SCHÖLL, M.; BAI, H.; WERNER, T. J.; ALAVI, A.; BUCHPIGUEL, C. Multimodality imaging approach in alzheimer disease. part i: structural mri, functional mri, diffusion tensor imaging and magnetization transfer imaging. *Dementia & Neuropsychologia*, SciELO Brasil, v. 9, n. 4, p. 318–329, 2015.

RAMÍREZ, J.; GÓRRIZ, J.; ORTIZ, A.; MARTÍNEZ-MURCIA, F.; SEGOVIA, F.; SALAS-GONZALEZ, D.; CASTILLO-BARNES, D.; ILLÁN, I.; PUNTONET, C.; INITIATIVE, A. D. N. et al. Ensemble of random forests one vs. rest classifiers for MCI and AD prediction using ANOVA cortical and subcortical feature selection and partial least squares. *Journal of Neuroscience Methods*, Elsevier, v. 302, p. 47–57, 2018.

RATTANABANNAKIT, C.; RISACHER, S. L.; GAO, S.; LANE, K. A.; BROWN, S. A.; MCDONALD, B. C.; UNVERZAGT, F. W.; APOSTOLOVA, L. G.; SAYKIN, A. J.; FARLOW, M. R. The cognitive change index as a measure of self and informant perception of cognitive decline: relation to neuropsychological tests. *Journal of Alzheimer's Disease*, IOS Press, v. 51, n. 4, p. 1145–1155, 2016.

RAZZAK, M. I.; NAZ, S.; ZAIB, A. Deep learning for medical image processing: Overview, challenges and the future. *Classification in BioApps*, Springer, p. 323–350, 2018.

REITAN, R. M. Validity of the trail making test as an indicator of organic brain damage. *Perceptual and Motor Skills*, SAGE Publications Sage CA: Los Angeles, CA, v. 8, n. 3, p. 271–276, 1958.

RISACHER, S. L.; KIM, S.; NHO, K.; FOROUD, T.; SHEN, L.; PETERSEN, R. C.; JR, C. R. J.; BECKETT, L. A.; AISEN, P. S.; KOEPPE, R. A. et al. APOE effect on Alzheimer's disease biomarkers in older adults with significant memory concern. *Alzheimer's & Dementia*, Wiley Online Library, v. 11, n. 12, p. 1417–1429, 2015.

RODRIGUES, A. L.; SANTANA, M. A. de; AZEVEDO, W. W.; BEZERRA, R. S.; BARBOSA, V. A.; LIMA, R. C. de; SANTOS, W. P. dos. Identification of mammary lesions in thermographic images: feature selection study using genetic algorithms and particle swarm optimization. *Research on Biomedical Engineering*, Springer, v. 35, n. 3, p. 213–222, 2019.

ROSEN, W.; MOHS, R.; DAVIS, K. A new rating scale for alzheimer's disease. *The American Journal of Psychiatry*, v. 141, n. 11, p. 1356–1364, 1984.

ROSEN, W. G.; TERRY, R. D.; FULD, P. A.; KATZMAN, R.; PECK, A. Pathological verification of ischemic score in differentiation of dementias. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, Wiley Online Library, v. 7, n. 5, p. 486–488, 1980.

ROWE, C. C.; ELLIS, K. A.; RIMAJOVA, M.; BOURGEAT, P.; PIKE, K. E.; JONES, G.; FRIPP, J.; TOCHON-DANGUY, H.; MORANDEAU, L.; O'KEEFE, G. et al. Amyloid imaging results from the Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging. *Neurobiology of Aging*, Elsevier, v. 31, n. 8, p. 1275–1283, 2010.

RUSSELL, S.; NORVIG, P. *Artificial Intelligence: A Modern Approach (3rd Edition)*. Pearson, 2009. ISBN 0136042597. Disponível em: <<https://www.amazon.com/Artificial-Intelligence-Modern-Approach-3rd/dp/0136042597?SubscriptionId=AKIAIOBINVZYXZQZ2U3A&tag=chimbori05-20&linkCode=xm2&camp=2025&creative=165953&creativeASIN=0136042597>>.

- SALVATORE, C.; CERASA, A.; CASTIGLIONI, I. MRI characterizes the progressive course of AD and predicts conversion to Alzheimer's dementia 24 months before probable diagnosis. *Frontiers in Aging Neuroscience*, v. 10, n. MAY, p. 1–13, 2018. ISSN 16634365.
- SAMPEDRO, F.; VILAPLANA, E.; LEON, M. J. D.; ALCOLEA, D.; PEGUEROLES, J.; MONTAL, V.; CARMONA-IRAGUI, M.; SALA, I.; SÁNCHEZ-SAUDINOS, M.-B.; ANTÓN-AGUIRRE, S. et al. APOE-by-sex interactions on brain structure and metabolism in healthy elderly controls. *Oncotarget*, Impact Journals, LLC, v. 6, n. 29, p. 26663, 2015.
- SANTANA, M.; PEREIRA, J. M. S.; LIMA, R.; SANTOS, W. Breast lesions classification in frontal thermographic images using intelligent systems and moments of haralick and zernike. In: *Understanding a cancer diagnosis*. [S.l.]: Nova Science, 2020. p. 65–80.
- SANTANA, M. A. d.; PEREIRA, J. M. S.; SILVA, F. L. d.; LIMA, N. M. d.; SOUSA, F. N. d.; ARRUDA, G. M. S. d.; LIMA, R. d. C. F. d.; SILVA, W. W. A. d.; SANTOS, W. P. d. Breast cancer diagnosis based on mammary thermography and extreme learning machines. *Research on Biomedical Engineering*, SciELO Brasil, v. 34, n. 1, p. 45–53, 2018.
- SANTANA, M. A. de; PEREIRA, J. M. S.; LIMA, C. L. de; ALMEIDA, M. B. J. de; ANDRADE, J. F. S. de; SOUZA, T. K. S. de; LIMA, R. d. C. F. de; SANTOS, W. P. dos. Classification of breast lesions in frontal thermographic images using a diagnosis aid intelligent system. In: *Biomedical Computing for Breast Cancer Detection and Diagnosis*. [S.l.]: IGI Global, 2021. p. 28–46.
- SANTOS, W. P. dos; ASSIS, F. M. de; SOUZA, R. E. de; FILHO, P. B. dos S. Dialectical classification of MR images for the evaluation of Alzheimer's disease. In: NAIK, G. R. (Ed.). *Recent Advances in Biomedical Engineering*. Vukovar: InTech, 2009. p. 211–222.
- SANTOS, W. P. dos; ASSIS, F. M. de; SOUZA, R. E. de; MENDES, P. B.; MONTEIRO, H. S. de S.; ALVES, H. D. A dialectical method to classify Alzheimer's magnetic resonance images. In: SANTOS, W. P. dos (Ed.). *Evolutionary Computation*. Vukovar: InTech, 2009. p. 473–486.
- SANTOS, W. P. dos; ASSIS, F. M. de; SOUZA, R. E. de; MENDES, P. B.; MONTEIRO, H. S.; ALVES, H. D. Dialectical non-supervised image classification. In: IEEE. *2009 IEEE Congress on Evolutionary Computation*. [S.l.], 2009. p. 2480–2487.
- SANTOS, W. P. dos; SOUZA, R. E. de; FILHO, P. B. S.; NETO, F. B. L.; ASSIS, F. M. de. A dialectical approach for classification of DW-MR Alzheimer's images. In: IEEE. *2008 IEEE Congress on Evolutionary Computation (IEEE World Congress on Computational Intelligence)*. [S.l.], 2008. p. 1728–1735.
- SCHMITZ, T. W.; SPRENG, R. N.; WEINER, M. W.; AISEN, P.; PETERSEN, R.; JACK, C. R.; JAGUST, W.; TROJANOWKI, J. Q.; TOGA, A. W.; BECKETT, L. et al. Basal forebrain degeneration precedes and predicts the cortical spread of alzheimer's pathology. *Nature communications*, Nature Publishing Group, v. 7, p. 13249, 2016.
- SCHUFF, N.; WOERNER, N.; BORETA, L.; KORNFIELD, T.; SHAW, L.; TROJANOWSKI, J.; THOMPSON, P.; JR, C. J.; WEINER, M.; INITIATIVE, A. D. N. Mri of hippocampal volume loss in early alzheimer's disease in relation to apoe genotype and biomarkers. *Brain*, Oxford University Press, v. 132, n. 4, p. 1067–1077, 2009.

- SHAW, L. M.; VANDERSTICHELE, H.; KNAPIK-CZAJKA, M.; FIGURSKI, M.; COART, E.; BLENNOW, K.; SOARES, H.; SIMON, A. J.; LEWCZUK, P.; DEAN, R. A. et al. Qualification of the analytical and clinical performance of csf biomarker analyses in adni. *Acta neuropathologica*, Springer, v. 121, n. 5, p. 597–609, 2011.
- SHEIKH, J. Geriatric depression scale (GDS): recent evidence and development of a shorter version. *Clin Gerontol*, v. 5, p. 165–173, 1986.
- SHI, F.; WANG, L.; DAI, Y.; GILMORE, J. H.; LIN, W.; SHEN, D. LABEL: pediatric brain extraction using learning-based meta-algorithm. *NeuroImage*, Elsevier, v. 62, n. 3, p. 1975–1986, 2012.
- SILVA, A. L. R. da; SANTANA, M. A. de; LIMA, C. L. de; ANDRADE, J. F. S. de; SOUZA, T. K. S. de; ALMEIDA, M. B. J. de; SILVA, W. W. A. da; LIMA, R. d. C. F. de; SANTOS, W. P. dos. Features selection study for breast cancer diagnosis using thermographic images, genetic algorithms, and particle swarm optimization. *International Journal of Artificial Intelligence and Machine Learning (IJAIML)*, IGI Global, v. 11, n. 2, p. 1–18, 2021.
- SILVA, I. R.; SILVA, G. S.; SOUZA, R. G. de; SANTOS, W. P. dos; ROBERTA, A. d. A. Model based on deep feature extraction for diagnosis of alzheimer's disease. In: IEEE. *2019 International Joint Conference on Neural Networks (IJCNN)*. [S.l.], 2019. p. 1–7.
- SILVA, I. R.; SOUZA, R. G. de; SILVA, G.; OLIVEIRA, C. S. de; CAVALCANTI, L. H.; BEZERRA, R. S.; ROBERTA, A. d. A.; SANTOS, W. P. dos. Utilização de redes convolucionais para classificação e diagnóstico da doença de Alzheimer. *II Simpósio de Inovação em Engenharia Biomédica*, p. 73–76, 2018.
- SILVA, I. R. R. da; SILVA, G. d. S. L. e; SOUZA, R. G. de; SANTANA, M. A. de; SILVA, W. W. A. da; LIMA, M. E. de; SOUZA, R. E. de; FAGUNDES, R.; SANTOS, W. P. dos. Deep learning for early diagnosis of Alzheimer's disease: a contribution and a brief review. In: *Deep Learning for Data Analytics*. [S.l.]: Elsevier, 2020. p. 63–78.
- SILVA, I. R. R. da; SILVA, G. dos Santos Lucas e; SOUZA, R. G. de; SANTANA, M. A. de; SILVA, W. W. A. da; LIMA, M. E.; SOUZA, R. E. de; FAGUNDES, R.; SANTOS, W. P. dos. Deep learning for early diagnosis of Alzheimer's disease: a contribution and a brief review. In: *Deep Learning for Data Analytics: Foundations, Biomedical Applications, and Challenges*. [S.l.]: Elsevier, 2020. p. 63–78.
- ŠIMIĆ, G.; LUCASSEN, P. J.; KRŠNIK, Ž.; KRUŠLIN, B.; KOSTOVIĆ, I.; WINBLAD, B.; BOGDANOVIĆ, N. nNOS expression in reactive astrocytes correlates with increased cell death related DNA damage in the hippocampus and entorhinal cortex in Alzheimer's disease. *Experimental Neurology*, Elsevier, v. 165, n. 1, p. 12–26, 2000.
- SMITH, S. M. Fast robust automated brain extraction. *Human Brain Mapping*, Wiley Online Library, v. 17, n. 3, p. 143–155, 2002.
- SØRENSEN, T. J. *A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons*. [S.l.]: Munksgaard Copenhagen, 1948. v. 5.
- SOUZA, R. G. de; SILVA, G. dos Santos Lucas e; SANTOS, W. P. dos; LIMA, M. E. de. Computer-aided diagnosis of alzheimer's disease by mri analysis and evolutionary computing. *Research on Biomedical Engineering*, Springer, v. 37, n. 3, p. 455–483, 2021.

- STEIN, D. J.; NASLUND, J. A.; BANTJES, J. COVID-19 and the global acceleration of digital psychiatry. *The Lancet Psychiatry*, Elsevier, v. 9, n. 1, p. 8–9, 2022.
- TOMBAUGH, T. N. Trail Making Test A and B: normative data stratified by age and education. *Archives of Clinical Neuropsychology*, Elsevier, v. 19, n. 2, p. 203–214, 2004.
- TONDELLI, M.; WILCOCK, G. K.; NICHELLI, P.; JAGER, C. A. D.; JENKINSON, M.; ZAMBONI, G. Structural MRI changes detectable up to ten years before clinical Alzheimer's disease. *Neurobiology of Aging*, Elsevier, v. 33, n. 4, p. 825–e25, 2012.
- TORGO, L.; RIBEIRO, R. P.; PFAHRINGER, B.; BRANCO, P. SMOTE for regression. In: SPRINGER. *Portuguese Conference on Artificial Intelligence*. [S.l.], 2013. p. 378–389.
- TOSUN, D.; JOSHI, S.; WEINER, M. W.; INITIATIVE, A. D. N. Neuroimaging predictors of brain amyloidosis in mild cognitive impairment. *Annals of Neurology*, Wiley Online Library, v. 74, n. 2, p. 188–198, 2013.
- TOSUN, D.; SCHUFF, N.; TRURAN-SACREY, D.; SHAW, L. M.; TROJANOWSKI, J. Q.; AISEN, P.; PETERSON, R.; WEINER, M. W.; INITIATIVE, A. D. N. et al. Relations between brain tissue loss, csf biomarkers, and the apoe genetic profile: a longitudinal mri study. *Neurobiology of aging*, Elsevier, v. 31, n. 8, p. 1340–1354, 2010.
- TSENG, J.; SAMAGH, S.; FRASER, D.; LANDMAN, A. B. Catalyzing healthcare transformation with digital health: Performance indicators and lessons learned from a digital health innovation group. In: ELSEVIER. *Healthcare*. [S.l.], 2018. v. 6, n. 2, p. 150–155.
- VAKIL, E.; BLACHSTEIN, H. Rey auditory-verbal learning test: Structure analysis. *Journal of Clinical Psychology*, Wiley Online Library, v. 49, n. 6, p. 883–890, 1993.
- VICHIANIN, Y.; KAREESAW, A.; CHAWALPARIT, O.; OHKI, M. Optimizing fractional intensity threshold for fsibrain extraction tool (bet) and comparing with freesurfer on 3d t1w mr images. *Siriraj Medical Journal*, v. 70, n. 5, p. 391–396, 2018.
- WANG, X.; LI, X.-H.; CHO, J. W.; RUSS, B. E.; RAJAMANI, N.; OMELCHENKO, A.; AI, L.; KORCHMAROS, A.; SAWIAK, S.; BENN, R. A. et al. U-net model for brain extraction: Trained on humans for transfer to non-human primates. *NeuroImage*, Elsevier, v. 235, p. 118001, 2021.
- WEINER, M. W.; AISEN, P. S.; JR, C. R. J.; JAGUST, W. J.; TROJANOWSKI, J. Q.; SHAW, L.; SAYKIN, A. J.; MORRIS, J. C.; CAIRNS, N.; BECKETT, L. A. et al. The alzheimer's disease neuroimaging initiative: progress report and future plans. *Alzheimer's & Dementia*, Wiley Online Library, v. 6, n. 3, p. 202–211, 2010.
- WEINER, M. W.; VEITCH, D. P.; AISEN, P. S.; BECKETT, L. A.; CAIRNS, N. J.; GREEN, R. C.; HARVEY, D.; JACK, C. R.; JAGUST, W.; LIU, E. et al. The alzheimer's disease neuroimaging initiative: a review of papers published since its inception. *Alzheimer's & Dementia*, Elsevier, v. 9, n. 5, p. e111–e194, 2013.
- WEINER, M. W.; VEITCH, D. P.; AISEN, P. S.; BECKETT, L. A.; CAIRNS, N. J.; CEDARBAUM, J.; DONOHUE, M. C.; GREEN, R. C.; HARVEY, D.; JR, C. R. J. et al. Impact of the alzheimer's disease neuroimaging initiative, 2004 to 2014. *Alzheimer's & Dementia*, Elsevier, v. 11, n. 7, p. 865–884, 2015.

WEINER, M. W.; VEITCH, D. P.; AISEN, P. S.; BECKETT, L. A.; CAIRNS, N. J.; GREEN, R. C.; HARVEY, D.; JR, C. R. J.; JAGUST, W.; MORRIS, J. C. et al. The Alzheimer's disease neuroimaging initiative 3: Continued innovation for clinical trial improvement. *Alzheimer's & Dementia*, Elsevier, v. 13, n. 5, p. 561–571, 2017.

WINBLAD, B.; PALMER, K.; KIVIPELTO, M.; JELIC, V.; FRATIGLIONI, L.; WAHLUND, L.-O.; NORDBERG, A.; BÄCKMAN, L.; ALBERT, M.; ALMKVIST, O. et al. Mild cognitive impairment—beyond controversies, towards a consensus: report of the international working group on mild cognitive impairment. *Journal of internal medicine*, Wiley Online Library, v. 256, n. 3, p. 240–246, 2004.

YOUNG, A. L.; OXTOBY, N. P.; DAGA, P.; CASH, D. M.; INITIATIVE on behalf of the A. D. N.; FOX, N. C.; OURSELIN, S.; SCHOTT, J. M.; ALEXANDER, D. C. A data-driven model of biomarker changes in sporadic alzheimer's disease. *Brain*, v. 137, n. 9, p. 2564–2577, 2014. Disponível em: <<http://dx.doi.org/10.1093/brain/awu176>>.

YUSHKEVICH, P. A.; AVANTS, B. B.; DAS, S. R.; PLUTA, J.; ALTINAY, M.; CRAIGE, C.; INITIATIVE, A. D. N. et al. Bias in estimation of hippocampal atrophy using deformation-based morphometry arises from asymmetric global normalization: an illustration in ADNI 3 T MRI data. *NeuroImage*, Elsevier, v. 50, n. 2, p. 434–445, 2010.

ZHANG, S.; SMILAGIC, N.; HYDE, C.; NOEL-STORR, A. H.; TAKWOINGI, Y.; MCSHANE, R.; FENG, J. 11 c-pib-pet for the early diagnosis of alzheimer's disease dementia and other dementias in people with mild cognitive impairment (mci). *Cochrane Database of Systematic Reviews*, John Wiley & Sons, Ltd, n. 7, 2014.