



**UNIVERSIDADE FEDERAL DO CEARÁ**  
**CENTRO DE TECNOLOGIA**  
**DEPARTAMENTO DE TELEINFORMÁTICA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TELEINFORMÁTICA**  
**DOUTORADO EM ENGENHARIA DE TELEINFORMÁTICA**

**ALDÍSIO GONÇALVES MEDEIROS**

**DESCRITOR DE CARACTERÍSTICAS EM COMPONENTES ESPECTRAIS**  
**BASEADO EM DIGRAFO PARA DETECÇÃO DE DOENÇAS**  
**NEURODEGENERATIVAS**

**FORTALEZA**

**2024**

ALDÍSIO GONÇALVES MEDEIROS

DESCRITOR DE CARACTERÍSTICAS EM COMPONENTES ESPECTRAIS BASEADO EM  
DIGRAFO PARA DETECÇÃO DE DOENÇAS NEURODEGENERATIVAS

Tese apresentada ao Programa de Pós-Graduação em Engenharia de Teleinformática do Centro de Tecnologia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de doutor em Engenharia de Teleinformática. Área de Concentração: Sinais e Sistemas

Orientador: Prof. Dr. Pedro Pedrosa Rebouças Filho

FORTALEZA

2024

Dados Internacionais de Catalogação na Publicação  
Universidade Federal do Ceará  
Sistema de Bibliotecas  
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

---

M438d Medeiros, Aldisio Gonçalves.

Descritor de características em componentes espectrais baseado em digrafo para detecção de doenças neurodegenerativas / Aldisio Gonçalves Medeiros. – 2024.  
155 f. : il. color.

Tese (doutorado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia de Teleinformática, Fortaleza, 2024.

Orientação: Prof. Dr. Pedro Pedrosa Rebouças Filho.

1. Extrator de atributos. 2. Inteligência artificial. 3. Deep learning. 4. Doenças neurodegenerativas. 5. Eletroencefalograma. I. Título.

CDD 621.38

---

ALDÍSIO GONÇALVES MEDEIROS

DESCRITOR DE CARACTERÍSTICAS EM COMPONENTES ESPECTRAIS BASEADO EM  
DIGRAFO PARA DETECÇÃO DE DOENÇAS NEURODEGENERATIVAS

Tese apresentada ao Programa de Pós-Graduação em Engenharia de Teleinformática do Centro de Tecnologia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de doutor em Engenharia de Teleinformática. Área de Concentração: Sinais e Sistemas

Aprovada em: 29 de Janeiro de 2024

BANCA EXAMINADORA

---

Prof. Dr. Pedro Pedrosa Rebouças  
Filho (Orientador)  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Plácido Rogério Pinheiro  
Universidade de Fortaleza (Unifor)

---

Prof. Dr. Carlos Maurício J. M. Dourado Jr  
Instituto Federal de Educação, Ciência e  
Tecnologia do Ceará (IFCE)

---

Prof. Dr. Fabrício Gonzalez Nogueira  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Julio César Santos dos Anjos  
Universidade Federal do Ceará (UFC)

Às minhas filhas, Clara Louise e Luna Liss.

## AGRADECIMENTOS

A Deus, por ter me guiado pelo caminho até aqui, ainda que com desafios, mas sempre mostrando os passos certos a serem seguidos;

Aos meus pais, Arilene Gonçalves e Aldízio de Souza, pelo amor e atenção dedicados ao longo da minha formação e construção como pessoa, além do apoio nos estudos sempre me guiaram no sentido de se tornar um ser humano melhor;

Às minhas filhas, Clara Louise e Luna Liss, o amor de vocês deram tons coloridos aos meus dias mais cinzas; o carinho e o sorriso de vocês trouxeram luz e alegria os meus dias mais sombrios, me fortaleceram quando em mim já não haviam forças, obrigado por me motivarem a concluir esta jornada;

À Samara Pinheiro, por todo apoio, incentivo e compreensão ofertados;

Ao Renan, obrigado pela força, mesmo distante, me fortalecia com seu encorajamento;

Ao meu orientador, Prof. Dr. Pedro Pedrosa Rebouças Filho, minha sincera gratidão por iluminar o caminho da minha trajetória acadêmica com sabedoria, serenidade, paciência e seriedade; além do companheirismo profissional, agradeço por poder chamá-lo de amigo;

Aos amigos de laboratório Solon, Lucas, Leandro, Hércules, Elene, e Navar, sempre presentes nos diferentes momentos desta jornada de pesquisa;

À Êmile, pela torcida, ideias e motivação que contribuíram muitíssimo para esta pesquisa;

À Natiele, pelo incentivo e pelo apoio em momentos turbulentos desta jornada;

À Flaviene Brito, pela dedicação e apoio nos meus momentos emocionais mais difíceis, obrigado pela sua condução durante este processo;

Aos amigos Valber e Hubson, pelas conversas e apoio ofertados;

Por fim, à todos que formam o Laboratório de Processamento de imagens, sinais e computação aplicada (LAPISCO), pelas calorosas discussões científico-filosóficas sempre muito produtivas e que contribuíram para a realização deste trabalho.

“Ninguém pode construir em teu lugar as pontes  
que precisarás passar, para atravessar o rio da  
vida – ninguém, exceto tu, só tu.”

(Friedrich Nietzsche)

## RESUMO

Esta pesquisa apresenta um método para descrição de sinais biomédicos com uma abordagem baseada na construção de um grafo direcionado, sugere também um processamento sobre os coeficientes espectrais do sinal. Tal processamento baseia-se nas transições dos coeficientes identificados quanto aos diferentes níveis de magnitude, por meio do marcador espectral escolhido. Neste contexto são propostos quatro diferentes marcadores espectrais que conduzem a análise das componentes de frequência e são baseados em diferentes pontos locais do sinal. A metodologia apresenta uma implementação simples e adaptativa, com capacidade de representar um ou mais sinais do mesmo paciente na mesma estrutura, sendo adequada para aplicações monocanal e multicanal. No contexto de doenças neurodegenerativas, foram escolhidas três aplicações dentre as patologias mais incidentes. Na primeira, o método extrai características de sinais de voz dos pacientes, para a tarefa de classificação dos saudáveis em relação aos portadores da Doença de Huntington; nessa aplicação foi utilizado um *dataset* com registros de voz. Além do teste estatístico baseado na análise de variância, também considerou-se medidas de acerto da matriz de confusão. Já a segunda aplicação avalia o método em um cenário multicanal, em especial para sinais de eletroencefalograma para a classificação de pacientes saudáveis em relação aos portadores da Doença de Alzheimer, foram utilizados cinco diferentes *datasets*, sendo os dois primeiros para classificação binária (saudáveis e doentes), já os três últimos apresentavam amostras com outras doenças neurodegenerativas e foram considerados para o cenário multiclasse. Por fim, a terceira aplicação investiga o desempenho do método na representação de pacientes com a Doença de Parkinson, neste cenário foram utilizados dois diferentes *datasets* já avaliados em estudos anteriores com pacientes saudáveis e doentes. Em todos os casos, como resultado do processamento do sinal a metodologia produz um tensor tridimensional, composto por três diferentes perspectivas do digrafo. O tensor espectral combinado com redes neurais profundas obteve resultados iguais ou superiores quando comparado com abordagens especialistas em cada aplicação, apresentando taxas de acerto superiores a 95% em todos os oito *datasets* considerados. Evidenciando que o grafo direcionado consegue capturar informações relevantes e a estrutura em tensor viabiliza uma combinação promissora com redes neurais profundas além de oferecer uma alternativa a extração de atributos associados às doenças neurodegenerativas.

**Palavras-chave:** descritor de características; EEG; voz; doenças neurodegenerativas; doença de Huntington; doença de Alzheimer; doença de Parkinson

## ABSTRACT

This research presents a method for describing biomedical signals based on constructing a directed graph, which also suggests processing the signal's spectral coefficients. This processing is based on the transitions of the coefficients identified in terms of the different levels of magnitude using the spectral marker chosen. In this context, four different spectral markers are proposed to analyze the frequency components and are based on different local points in the signal. The methodology has a simple and adaptive implementation. It can represent one or more signals from the same patient in the same structure, making it suitable for single-channel and multichannel applications. In the context of neurodegenerative diseases, three applications were chosen from among the most common pathologies. At first, the method extracts characteristics from patients' voice signals to classify healthy patients in relation to those with Huntington's disease; in this application, a dataset with voice records was used. In addition to the statistical test based on analysis of variance, measures of the correctness of the confusion matrix were also considered. The second application evaluates the method in a multichannel scenario, particularly for electroencephalogram signals for classifying healthy patients about those with Alzheimer's disease. Five different datasets were used, the first two for binary classification (healthy and sick). At the same time, the last three had samples with other neurodegenerative diseases and were considered for the multiclass scenario. Finally, the third application investigates the method's performance in representing patients with Parkinson's disease. In this scenario, two datasets were used that had already been evaluated in previous studies with healthy and sick patients. In all cases, as a result of signal processing, the methodology produces a three-dimensional tensor made up of three different perspectives of the digraph. The spectral tensor combined with deep neural networks obtained equal or better results when compared to expert approaches in each application, with hit rates above 95% in all eight datasets considered. This shows that the directed graph can capture relevant information, and the tensor structure makes a promising combination with deep neural networks feasible and offers an alternative to extracting attributes associated with neurodegenerative diseases.

**Keywords:** features descriptor; EEG; voice; neurodegenerative diseases; Huntington's disease; Alzheimer's disease; Parkinson's disease

## LISTA DE FIGURAS

Figura 1 – Representação do sinal sob duas perspectivas: domínio do tempo (em laranja), com suas respectivas frequências decompostas (em outras cores); e no domínio da frequência, componentes espectrais representados por barras (em azul) . . . . .	40
Figura 2 – Sinais EEG perturbados por ruído e artefatos e sua potência espectral correspondente. (a) sinal EEG sem ruídos registrado no eletrodo T3. (b) sinal EEG, registrado no eletrodo O1, perturbado pelo ruído da linha de energia. (c) Sinal registrado no eletrodo Fp1 contendo movimento ocular (esquerda) e piscar de olhos artefatos (à direita). (d) Sinal registrado no eletrodo T3, contendo artefato de movimento muscular. . . . .	42
Figura 3 – Função de custo $J$ de $w$ apresentando a descida do gradiente e taxa de ajuste do vetor $W$ . Em verde, a configuração de pesos iniciais, em laranja, a configuração desejada. . . . .	54
Figura 4 – Bloco Residual. . . . .	55
Figura 5 – Operação de convolução em um sinal bidimensional $S$ sob um filtro convolucional $K$ . . . . .	57
Figura 6 – Matriz de confusão, apresentando os diferentes valores de verdadeiros positivos, falsos positivos, falsos negativos e verdadeiros negativos. . . . .	61
Figura 7 – Exemplo visual da metodologia de validação cruzada com amostragem estratificada. . . . .	62
Figura 8 – Fluxo da metodologia de extração proposta. . . . .	65
Figura 9 – Representação do sinal sob o domínio do tempo e frequência. . . . .	66
Figura 10 – Análise proposta dos extremos locais relativos da função $\phi$ . . . . .	67
Figura 11 – Marcadores espectrais alfa ( $\alpha$ ) e beta ( $\beta$ ) propostos no sinal da função $\phi$ . . . . .	68
Figura 12 – Marcadores espectrais gamma ( $\gamma$ ) e delta ( $\delta$ ) propostos no sinal da função $\phi$ . . . . .	70
Figura 13 – Amplitude máxima e mínima dos coeficientes e a divisão em níveis de magnitude. . . . .	72
Figura 14 – Construção do mapa de transições do marcador no domínio da frequência. . . . .	73
Figura 15 – Fluxo de processamento para cenário multicanal. . . . .	75
Figura 16 – Grafo direcionado $DG_1$ construído a partir do marcador $\alpha$ com $n(V) = 5$ e $n(E) = 6$ . . . . .	76

Figura 17 – Grafo direcionado $DG_2$ construído a partir do marcador $\beta$ com $n(V) = 5$ e $n(E) = 6$ . . . . .	76
Figura 18 – Grafo direcionado $DG_3$ construído a partir do marcador $\gamma$ com $n(V) = 9$ e $n(E) = 13$ . . . . .	77
Figura 19 – Grafo direcionado $DG_4$ construído a partir do marcador $\delta$ com $n(V) = 9$ e $n(E) = 16$ . . . . .	78
Figura 20 – A soma das arestas incidentes no vértice denota o quantitativo de coeficientes de magnitude presentes no respectivo nível representado. . . . .	79
Figura 21 – A soma das arestas incidentes no vértice denota o quantitativo de coeficientes de magnitude presentes no nível representado pelo vértice. . . . .	79
Figura 22 – Junção das diferentes perspectivas do grafo representadas por cada matriz em uma estrutura tensorial denominada tensor espectral. . . . .	82
Figura 23 – Amostra do sinal sintético formado por 5 frequências com magnitudes distintas. . . . .	90
Figura 24 – Amostra do sinal sintético formado por 10 frequências com magnitudes distintas. . . . .	91
Figura 25 – Amostras sintéticas no domínio da frequência, em azul o sinal formado por 5 frequências, já em vermelho o sinal formado por 10 frequências distintas. . . . .	91
Figura 26 – Exemplo da divisão em níveis de magnitude do sinal sintético da classe 1, variando $L$ de 3 até 15. Cada nível é representando por um vértices no digrafo. . . . .	92
Figura 27 – Exemplo de grafos produzidos a partir do mesmo sinal no domínio do tempo (amostras à esquerda) e no domínio da frequência (amostras à direita). Nota-se uma simplificação da representação comparando as duas perspectivas do mesmo sinal. . . . .	93
Figura 28 – Exemplo de digrafos criados a partir de duas amostras, uma do sinal da classe 0 (em azul) e uma do sinal da classe 1 (em vermelho). Ambas sob diferentes construções com $L$ assumindo os valores 3, 6 e 12. . . . .	94
Figura 29 – Histograma das amostras quanto a sua distribuição e média das classes, de acordo com a extração de atributos representada pelas matrizes de representação do grafo. . . . .	96
Figura 30 – Resultado do valor- $p$ no teste ANOVA para discriminação entre as duas classes de sinais. Em cada matriz, as linhas representam o número de vértices e as colunas, mas matrizes de representação do grafo. . . . .	97

Figura 31 – Principais degradações no corpo estriado dorsal do cérebro que afetam o controle parcial do movimento voluntário. Adaptado de Cervato (2023). . . . .	101
Figura 32 – Metodologia para a extração de características de sinais de voz, passos A-D. Nesta proposta, os pacientes são representados pelo tensor espectral do sinal.	104
Figura 33 – Resultado do valor- $p$ no teste ANOVA para discriminação entre as duas classes de sinais. Em cada matriz, as linhas representam o número de vértices e as colunas, mas matrizes de representação do grafo. . . . .	105
Figura 34 – Taxas de acerto de acordo com o número de nós para os marcadores espectrais	107
Figura 35 – Top 4 melhores modelos de classificação com base nas taxas de acerto . . . . .	108
Figura 36 – Desempenho dos marcadores $\alpha$ , $\beta$ , $\gamma$ , $\delta$ de acordo com o índice médio de Acurácia no conjunto de teste. . . . .	109
Figura 37 – Resultados do extratores clássicos e baseados em wavelets. . . . .	112
Figura 38 – Resultados em relação aos extratores de recursos de áudio. . . . .	114
Figura 39 – Atrofia cerebral devido ao acúmulo anormal de proteínas, como as placas de beta-amiloide e os emaranhados neurofibrilares na doença de Alzheimer e afeta regiões responsáveis pela memória, atrofia do Hipocampo, e capacidade de aprendizado e raciocínio, atrofia do córtex cerebral. Adaptado de Hosseini <i>et al.</i> (2020). . . . .	117
Figura 40 – Metodologia para extração dos diferentes sinais de eletroencefalograma. Para cada eletrodo, um grafo direcionado de arestas ponderadas é desenvolvido e representado como tensor para classificação via rede neural residual. . . . .	121
Figura 41 – Histograma das amostras extraídas para as classes AD vs. HC do <i>dataset I</i> . (L=12) . . . . .	124
Figura 42 – Histograma das amostras extraídas para as classes AD vs. HC do <i>dataset II</i> . (L=12) . . . . .	125
Figura 43 – Resultado do valor- $p$ no teste ANOVA para discriminação no <i>dataset I</i> (AD vs HC) . . . . .	126
Figura 44 – Resultado do valor- $p$ no teste ANOVA para discriminação no <i>dataset II</i> (AD vs HC) . . . . .	127
Figura 45 – Resultado do valor- $p$ no teste ANOVA para discriminação no <i>dataset III</i> (AD vs HC) . . . . .	127

Figura 46 – Resultado do valor- $p$ no teste ANOVA para discriminação de amostras no <i>dataset IV</i> com 2 classes (AD vs MCI) . . . . .	128
Figura 47 – Resultado do valor- $p$ no teste ANOVA para discriminação de amostras no <i>dataset V</i> (AD vs HC) . . . . .	128
Figura 48 – Histograma da extração para as classes AD vs. FTD vs. HC do <i>dataset III</i> . (L=30) . . . . .	129
Figura 49 – Histograma da extração para as classes AD vs. MCI vs. HC do <i>dataset V</i> . (L=6)	130
Figura 50 – Resultado do valor- $p$ no <i>dataset III</i> com 3 classes (AD vs FTD vs HC) . . .	130
Figura 51 – Resultado do valor- $p$ no <i>dataset V</i> com 3 classes (AD vs MCI vs HC) . . .	131
Figura 52 – Estado da doença de Parkinson ilustrando a degradação da substância negra, associada ao controle de movimentos. Adaptado de Polat e Nour (2020). . .	136
Figura 53 – Histograma das amostras extraídas para as classes PD vs. HC do <i>dataset I</i> . (L=12) . . . . .	140
Figura 55 – Resultado do valor- $p$ no teste ANOVA para discriminação no <i>dataset I</i> (PD vs HC) . . . . .	140
Figura 54 – Histograma das amostras extraídas para as classes PD vs. HC do <i>dataset II</i> . (L=12) . . . . .	141
Figura 56 – Resultado do valor- $p$ no teste ANOVA para discriminação no <i>dataset II</i> (PD vs HC) . . . . .	141

## LISTA DE TABELAS

Tabela 1 – Extratores de atributos clássicos e número total de atributos com base nos experimentos sobre amostras de voz de pacientes saudáveis e com a doença de Huntington. . . . .	45
Tabela 2 – Diferentes configurações de extratores produzidos pelo OpenSMILE e o número total de atributos. . . . .	48
Tabela 3 – Relação entre marcadores, número de vértices e medidas de distância entre as classes considerando o tensor espectral proposto. . . . .	99
Tabela 4 – Tempo médio de extração para 400 sinais com duração de 5 segundos, composto por diferentes frequências e com ruído induzido. . . . .	99
Tabela 5 – Lista dos diferentes modelos de classificação e seus respectivos espaços de parâmetros avaliados para otimização. Os parâmetros referem-se à terminologia do framework. . . . .	110
Tabela 6 – Comparação entre extratores clássicos na análise de sinal com a abordagem proposta . . . . .	111
Tabela 7 – Comparação entre extratores especializados em análise de áudio em relação à abordagem proposta . . . . .	113
Tabela 8 – Comparativamente com abordagens da literatura recente, os trabalhos abaixo visam classificar pacientes com doença de Huntington utilizando sinais de voz. . . . .	115
Tabela 9 – Resumo dos <i>datasets</i> I e II (Doença de Alzheimer vs. Saudável). . . . .	120
Tabela 10 – Resumo dos <i>datasets</i> III, IV e V (Doença de Alzheimer vs. Outras patologias). . . . .	120
Tabela 11 – Resultados para o melhor desempenho do método proposto nas amostras do dataset I e II, para classificação do Alzheimer em relação a pacientes saudáveis. . . . .	132
Tabela 12 – Resultados para o melhor desempenho do método proposto nas amostras do dataset III para classificação do Alzheimer em relação a pacientes com outras patologias. . . . .	133
Tabela 13 – Resultados para o melhor desempenho do método proposto nas amostras do dataset IV para classificação do Alzheimer em relação a pacientes com outras patologias. . . . .	133

Tabela 14 – Resultados para o melhor desempenho do método proposto nas amostras do dataset V para classificação do Alzheimer em relação a pacientes com outras patologias. . . . .	134
Tabela 15 – Sumário comparativo com trabalhos relacionados a classificação de Alzheimer. . . . .	134
Tabela 16 – Resultados para o melhor desempenho do método proposto nas amostras do dataset I e II, para classificação do Parkinson em relação a pacientes saudáveis.	143
Tabela 17 – Sumário comparativo com trabalhos relacionados a classificação de Parkinson.	143

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>19</b>
<b>1.1</b>	<b>Contextualização</b>	<b>19</b>
<b>1.2</b>	<b>Motivação</b>	<b>22</b>
<b>1.3</b>	<b>Trabalhos relacionados</b>	<b>23</b>
<b>1.3.1</b>	<b><i>Metodologias em monocanal</i></b>	<b>24</b>
<b>1.3.1.1</b>	<b><i>Abordagens relacionadas a Doença de Huntington</i></b>	<b>24</b>
<b>1.3.2</b>	<b><i>Metodologias multicanais</i></b>	<b>25</b>
<b>1.3.2.1</b>	<b><i>Abordagens relacionadas a Doença de Alzheimer</i></b>	<b>25</b>
<b>1.3.2.2</b>	<b><i>Abordagens relacionadas a Doença de Parkinson</i></b>	<b>27</b>
<b>1.4</b>	<b>Objetivos</b>	<b>29</b>
<b>1.4.1</b>	<b><i>Objetivos Específicos</i></b>	<b>29</b>
<b>1.5</b>	<b>Produção científica</b>	<b>30</b>
<b>1.6</b>	<b>Organização da Tese</b>	<b>32</b>
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>33</b>
<b>2.1</b>	<b>Teoria dos Grafos</b>	<b>33</b>
<b>2.1.1</b>	<b><i>Grafos</i></b>	<b>33</b>
<b>2.1.2</b>	<b><i>Digrafos</i></b>	<b>34</b>
<b>2.1.3</b>	<b><i>Matriz de Adjacência</i></b>	<b>36</b>
<b>2.1.4</b>	<b><i>Matriz Laplaciana</i></b>	<b>36</b>
<b>2.1.5</b>	<b><i>Matriz Laplaciana Combinatória</i></b>	<b>37</b>
<b>2.1.6</b>	<b><i>Matriz Laplaciana Combinatorial Direcionada</i></b>	<b>38</b>
<b>2.2</b>	<b>Transformada Discreta de Fourier (DFT)</b>	<b>39</b>
<b>2.2.1</b>	<b><i>Definição Matemática da DFT:</i></b>	<b>39</b>
<b>2.2.2</b>	<b><i>Relação da DFT com Ruídos nos Sinais</i></b>	<b>41</b>
<b>2.3</b>	<b>Extração clássica de atributos em sinal unidimensional</b>	<b>42</b>
<b>2.3.1</b>	<b><i>Extratores baseados em estatísticas do sinal:</i></b>	<b>43</b>
<b>2.3.2</b>	<b><i>Extratores baseados em wavelets</i></b>	<b>44</b>
<b>2.3.3</b>	<b><i>Extratores específicos para sinais de voz</i></b>	<b>46</b>
<b>2.4</b>	<b>Modelos de aprendizado de máquina: abordagem clássica</b>	<b>48</b>
<b>2.5</b>	<b>Redes Neurais Artificiais</b>	<b>52</b>

<b>2.5.1</b>	<b><i>Multilayer Perceptron (MLP)</i></b> . . . . .	52
2.5.1.1	<i>O problema do Vanishing Gradient</i> . . . . .	54
<b>2.5.2</b>	<b><i>Rede neural residual (ResNet)</i></b> . . . . .	54
<b>2.5.3</b>	<b><i>Redes Neurais Convolucionais</i></b> . . . . .	56
2.5.3.1	<i>Operação de Convolução</i> . . . . .	56
2.5.3.2	<i>Arquitetura MnasNet:</i> . . . . .	57
2.5.3.3	<i>Arquitetura MobileNet V2:</i> . . . . .	58
2.5.3.4	<i>Arquitetura ResNeXt:</i> . . . . .	58
2.5.3.5	<i>Arquitetura ShuffleNet V2:</i> . . . . .	58
2.5.3.6	<i>Arquitetura Wide ResNet:</i> . . . . .	59
<b>2.6</b>	<b><i>Avaliação de desempenho</i></b> . . . . .	59
<b>2.6.1</b>	<b><i>Análise de Variância (ANOVA) quanto a discriminação de classes</i></b> . . . . .	59
2.6.1.1	<i>ANOVA e o Valor-p para algoritmo descritores de características</i> . . . . .	60
<b>2.6.2</b>	<b><i>Métricas quanto ao Desempenho em Modelos de Classificação</i></b> . . . . .	60
<b>2.6.3</b>	<b><i>Conclusão do Capítulo</i></b> . . . . .	63
<b>3</b>	<b>DESCRITOR DE CARACTERÍSTICAS EM COMPONENTES ESPECTRAIS BASEADO EM DIGRAFO: A METODOLOGIA</b> . . . . .	64
<b>3.1</b>	<b>Domínio da frequência</b> . . . . .	64
<b>3.2</b>	<b>Encontrando pontos marcadores nas componentes espectrais do sinal</b> . . . . .	65
<b>3.2.1</b>	<b><i>Análise proposta dos extremos locais relativos nas componentes espectrais da função <math>\phi</math></i></b> . . . . .	66
<b>3.2.2</b>	<b><i>Ponto de inflexão nas componentes espectrais da função <math>\phi</math></i></b> . . . . .	67
3.2.2.1	<i>O marcador espectral alpha (<math>\alpha</math>) proposto</i> . . . . .	68
3.2.2.2	<i>O marcador espectral beta (<math>\beta</math>) proposto</i> . . . . .	69
3.2.2.3	<i>O marcador espectral gamma (<math>\gamma</math>) proposto</i> . . . . .	69
3.2.2.4	<i>O marcador espectral delta (<math>\delta</math>) proposto</i> . . . . .	70
<b>3.3</b>	<b>O digrafo como proposta de representação do sinal espectral</b> . . . . .	71
<b>3.3.1</b>	<b><i>O processamento multicanal</i></b> . . . . .	73
<b>3.3.2</b>	<b><i>Propriedades da abordagem proposta</i></b> . . . . .	74
3.3.2.1	<i>A proposta da composição em tensor tridimensional de componentes espectrais</i> . . . . .	79
<b>3.4</b>	<b>Algoritmo do descritor de sinais proposto baseado em grafo direcionado</b> . . . . .	82
<b>3.4.1</b>	<b><i>Complexidade computacional do algoritmo proposto</i></b> . . . . .	85

3.5	Algoritmo proposto para seleção dos parâmetros . . . . .	86
3.5.1	<i>Complexidade computacional do algoritmo proposto</i> . . . . .	87
3.6	Conclusão do Capítulo . . . . .	88
4	<b>RESULTADOS NO CONTEXTO DE AMOSTRAS SINTÉTICAS</b> . . .	89
4.1	Amostras sintéticas . . . . .	89
4.2	Construção do digrafo como descritor do sinal . . . . .	91
4.3	Análise da discriminação de amostras entre classes. . . . .	94
4.3.1	<i>Teste estatístico ANOVA</i> . . . . .	95
4.3.2	<i>Tempo de extração</i> . . . . .	98
4.4	Conclusão do Capítulo . . . . .	99
5	<b>RESULTADOS EM CONTEXTO MONOCANAL: CLASSIFICAÇÃO DA DOENÇA DE HUNTINGTON VIA SINAIS DE VOZ.</b> . . . . .	101
5.1	A doença de Huntington . . . . .	101
5.2	Trabalhos relacionados a detecção da doença . . . . .	102
5.3	Amostras de pacientes . . . . .	103
5.4	Metodologia . . . . .	103
5.4.1	<i>Construção do Digrafo</i> . . . . .	103
5.4.2	<i>Avaliação de Desempenho</i> . . . . .	103
5.5	Resultados e discussão . . . . .	105
5.5.1	<i>Teste estatístico ANOVA</i> . . . . .	105
5.5.2	<i>Comparação de marcadores de componentes espectrais quanto a classificação</i> . . . . .	106
5.5.3	<i>Comparação com extratores de atributos clássicos e wavelets</i> . . . . .	110
5.5.4	<i>Comparação com extratores de áudio</i> . . . . .	113
5.6	Conclusões do capítulo . . . . .	114
6	<b>RESULTADOS EM CONTEXTO MULTICANAL: CLASSIFICAÇÃO DA DOENÇA DE ALZHEIMER COM MÚLTIPLOS CANAIS DE EEG.</b> . . . . .	116
6.1	A doença de Alzheimer . . . . .	116
6.2	Trabalhos relacionados a detecção da doença . . . . .	117
6.3	Amostras de pacientes . . . . .	117
6.3.1	<i>Amostras de pacientes: Doença de Alzheimer vs. Saudável</i> . . . . .	118
6.3.1.1	<i>Dataset I</i> . . . . .	118
6.3.1.2	<i>Dataset II</i> . . . . .	118

6.3.2	<i>Amostras de pacientes: Doença de Alzheimer vs. Outras</i>	119
6.3.2.1	<i>Dataset III</i>	119
6.3.2.2	<i>Dataset IV</i>	119
6.3.2.3	<i>Dataset V</i>	119
6.4	<b>Metodologia</b>	121
6.4.1	<i>Avaliação de Desempenho</i>	123
6.5	<b>Resultados e discussão</b>	123
6.5.1	<i>Teste estatístico ANOVA</i>	123
6.5.2	<i>Teste com duas classes</i>	124
6.5.3	<i>Teste com três classes</i>	128
6.6	<b>Combinação dos marcadores das componentes espectrais com modelos de classificação</b>	131
6.7	<b>Conclusões do capítulo</b>	134
7	<b>RESULTADOS EM CONTEXTO MULTICANAL: CLASSIFICAÇÃO DA DOENÇA DE PARKINSON</b>	136
7.1	<b>A doença de Parkinson</b>	136
7.2	<b>Trabalhos Relacionados</b>	137
7.3	<b>Amostras de pacientes</b>	137
7.3.1	<i>Dataset I</i>	137
7.3.2	<i>Dataset II</i>	138
7.4	<b>Metodologia</b>	138
7.4.1	<i>Avaliação de Desempenho</i>	139
7.5	<b>Resultados e discussão</b>	139
7.5.1	<i>Teste estatístico ANOVA</i>	139
7.6	<b>Conclusões do Capítulo</b>	143
8	<b>CONCLUSÕES E TRABALHOS FUTUROS</b>	144
8.0.1	<i>Trabalhos Futuros</i>	146
	<b>REFERÊNCIAS</b>	147

# 1 INTRODUÇÃO

Esta tese propõe um novo métodos de extração de características relevantes em sinais unidimensionais biológicos, abrangendo uma estrutura baseada em grafo para aplicações monocanal e multicanal, o que torna a proposta flexível para contextos onde o número de sinais pode ser variável. Neste capítulo, apresentamos a contextualização da área, além da motivação e os objetivos da tese, bem como as principais contribuições e a produção científica.

## 1.1 Contextualização

Apesar de toda a beleza no instigante funcionamento do sistema nervoso central, onde aglomerados de células neurais operam em sintonia, promovendo sentimentos, percepções dos sentidos, abrangendo emoções associadas a memórias e a cognição, estão as doenças neurodegenerativas que comprometem o bom funcionamento de toda esta nobre complexidade do tecido neural. Do ponto de vista patológico, Appakaya *et al.* (2023) descreve que essas enfermidades manifestam-se por meio de agregados proteicos irregulares, e incluem distúrbios sinápticos e morte neuronal.

Dentre as mais incidentes está a Doença de Alzheimer, uma das formas mais prevalentes de demência e caracteriza-se pelo acúmulo de placas de beta-amiloide<sup>1</sup>, proteína que se acumula de modo anormal em regiões associadas a memória e cognição (MOOTOO *et al.*, 2023). Já no espectro do movimento, a Doença de Parkinson (PD) revela-se pela degeneração dopaminérgica na substância negra<sup>2</sup>, que está diretamente ligada ao controle de movimentos voluntários (APPAKAYA *et al.*, 2023), enquanto a Doença de Huntington (HD) origina-se de uma mutação no gene HTT<sup>3</sup>, desencadeando a produção anômala da proteína huntingtina (SUBRAMANIAN *et al.*, 2023), produzindo movimentos involuntários que afetam a locomoção e a voz.

O impacto social dessas condições é amplo e com aspectos multidisciplinares. Do ponto de vista econômico, há um ônus financeiro nos sistemas de saúde. Estudos como o de Wimo *et al.* (2017) apontam para um aumento exponencial nos custos associados ao cuidado de

---

<sup>1</sup> É uma proteína naturalmente presente no cérebro, com tendência para formar aglomerados. Quando associada ao receptor sináptico PirB, enfraquece a comunicação entre neurônios.

<sup>2</sup> Responsável por produzir o neurotransmissor dopamina, que é essencial para o controle motor e execução de funções cognitivas.

<sup>3</sup> Gene responsável pela codificação da proteína Huntingtina, quando mutante, sinaliza a produção anormal desta proteína, causando a morte de neurônios.

pacientes com demência, com custos globais da ordem de US\$ 1 trilhão. Mais recentemente, Porsteinsson *et al.* (2021) aponta que os custos relacionados à doença de Alzheimer ultrapassaram os 300 bilhões de dólares em 2021, já Dorsey *et al.* (2018) indica que mais de 52 bilhões são investidos anualmente com o tratamento da Doença de Parkinson, somente nos Estados Unidos.

No Brasil, estima-se que mais de 28 mil reais por ano são empregados no tratamento individual da Doença de Parkinson, sem contar o custo associado aos cuidadores como discute Bovolenta *et al.* (2017). Por outro lado, Veras *et al.* (2008) examinaram 41 famílias de portadores de demência no Rio de Janeiro e observaram que os custos associados ao tratamento representavam aproximadamente dois terços da renda familiar. Esse custo aumentava para 75% nos estágios iniciais da doença e alcançava 80% quando outras doenças crônicas são associadas.

Esses custos abrangem não apenas despesas diretas com cuidados médicos, mas também custos indiretos, como a perda de produtividade de pacientes e cuidadores. Há também impactos sociais, já que o declínio cognitivo e motor exige uma crescente dependência para as atividades diárias, sobrecarregando os cuidadores. Gauthier *et al.* (2013) argumenta que esse ônus não apenas compromete a qualidade de vida dos cuidadores, mas impacta na coesão familiar e nas relações interpessoais. Tal cenário, evidencia a necessidade de políticas públicas que incentivem o diagnóstico precoce, visando melhorias na sobrevivência e nos custos envolvidos.

A pesquisa tem avançado no sentido do diagnóstico automatizado e novos estudos emergem como uma necessidade imperativa. É na convergência da neurociência e da inteligência artificial (IA) que se vislumbra um horizonte promissor. Algoritmos de aprendizado profundo, treinados em grandes conjuntos de dados, apresentam uma precisão crescente no diagnóstico precoce, conforme evidenciado por estudos como o de Puri *et al.* (2023) e Appakaya *et al.* (2023) para a detecção de doenças neuropsiquiátricas.

Nesse campo, encontramos os exames de EEG como ferramenta valiosa no diagnóstico precoce, oferecendo análises fundamentais sobre as alterações elétricas no cérebro. Stam *et al.* (2007) destacam que os pacientes frequentemente apresentam padrões de atividade elétrica cerebral distintos, refletindo disfunções específicas relacionadas a essas condições. Parihar e Swami (2023) ressaltam a sensibilidade do EEG na detecção precoce de anomalias elétricas, permitindo intervenções terapêuticas em estágios iniciais. Já Gauthier *et al.* (2013), reforça como algoritmos de reconhecimento de padrões aplicados a sinais de EEG podem diferenciar padrões sutis, permitindo uma avaliação mais abrangente da progressão da doença.

Sob outra perspectiva, a análise dos sinais de voz tem se revelado uma ferramenta

promissora para o diagnóstico precoce de doenças neurodegenerativas, oferecendo uma abordagem não invasiva e de baixo custo. Estudos como os de Liu *et al.* (2023) destacam como pacientes afetados pelas doenças de Parkinson e Alzheimer apresentam mudanças distintas nos padrões acústicos e articulatórios da fala. Thies *et al.* (2020) ressaltam as vantagens da análise de sinais de voz para discernir padrões associados ao diagnóstico de diferentes estágios dessas doenças, proporcionando assim uma ferramenta alternativa, especialmente útil para o uso em populações carentes, onde o acesso a equipamentos sofisticados é impraticável.

Neste sentido, o diagnóstico automatizado das doenças neurodegenerativas depende substancialmente da extração de características e identificação de padrões discriminativos nos sinais. A seleção cuidadosa de características possibilita capturar nuances e variações não perceptíveis em uma análise direta dos dados brutos. Tais atributos representam propriedades estatísticas ou características específicas da doença, que são essenciais para representar as amostras e contribuir para a triagem de pacientes doentes dos saudáveis (LIU *et al.*, 2023). Os estudos de Parihar e Swami (2023) destacam que a extração e seleção de atributos contribui para a capacidade de reduzir a dimensionalidade dos dados, preservando, ao mesmo tempo, as informações críticas para a tomada de decisões, característica importante para sinais biológicos.

Diante deste cenário, esta tese concentra-se no desenvolvimento de um novo descritor de sinais. Inspirado na teoria dos grafos, a técnica abrange a análise de sinais sob a perspectiva de suas componentes espectrais, com ênfase nos coeficientes de magnitude que são representados pelos marcadores espectrais propostos. O grafo direcionado, também chamado de digrafo, possui arestas ponderadas, cujos vértices atuam como representação dos níveis de magnitude e suas relações com os coeficientes espectrais das frequências. A modelagem possibilita operações sobre a topologia do grafo, como cálculo de centralidade, cálculo de caminhos entre vértices, além de métricas de conectividade. Tal estrutura mostrou-se flexível, sendo aplicável a contexto monocanal e multicanal da aquisição dos sinais biológicos.

A proposta também sugere uma representação das amostras em uma estrutura tensorial, composta por três diferentes perspectivas extraídas do digrafo. Essa proposta inova trazendo um extrator compatível com redes neurais profundas, tipicamente adaptadas para classificação de tensores. Além disso, o uso de tensor abre margem para associação com tecnologias recentes como a transferência de aprendizado para modelos que tem como entrada tensores de três dimensões. Neste sentido, o estudo visa contribuir, na tentativa de expandir os horizontes para análise dos sinais de pacientes e fomentar o campo da classificação automática entre indivíduos

saudáveis e aqueles acometido por distúrbios neurodegenerativos.

## 1.2 Motivação

Segundo dados da Organização Mundial da Saúde (OMS), atualmente, mais de 55 milhões de pessoas sofrem de demência em todo o mundo. Dentre esse pacientes, são estimados que 60% vivem em países de baixa e média renda per capita. Essa enfermidade apresenta incidência crescente, anualmente são registrados aproximadamente 10 milhões de novos casos. E atua como a sétima causa quanto a mortalidade no mundo, dentre esses, estima-se que 65% são idosos do sexo feminino com mortes relacionadas a demência. A OMS indica que a causa mais comum de demência é a Doença de Alzheimer, que representa cerca de 60 a 80% dos idosos com demência, tendo também como causa a Doenças de Parkinson, em seus estágios mais avançados e a Doença de Huntington (WORLD HEALTH ORGANIZATION, 2022; HUANG, 2023).

No Brasil, apesar dos dados epidemiológicos serem escassos, uma vez que estas doenças ainda não fazem parte do Sistema de Doenças de Notificação Compulsória, pesquisas nacionais apresentam índices preocupantes. Segundo Jesus (2022), as estimativas indicam que há uma prevalência média de doenças neurodegenerativas no país com crescimento médio mais alto que a média mundial, especialmente na população com 65 anos ou mais, passando de 7,6% para 7,9% entre 2010 e 2020, ou seja, 55.000 novos casos por ano. Jesus (2022) argumenta que as pesquisas fazem estimativas com base nas concessões de auxílio-doença da previdência social, e foram estimados 13.899 auxílios-doença somente no período entre 2008 e 2017, o que representa 36,98% de novos casos no país. Tais levantamentos indicam que estudos brasileiros nesta área contribuem com o levantamento de dados para o fomento e incentivo de pesquisas nacionais nessa área.

Por outro lado, Huang (2023) relata que as doenças neurodegenerativas como doença de Parkinson e doença de Huntington são classificadas como doenças do movimento, caracterizando-se por lesões progressivas nos gânglios basais, agrupamentos de células diretamente relacionadas ao controle do movimento, que comprometem os movimentos voluntários e involuntários. Esse quadro debilita gradualmente a musculatura e o controle motor, colocando em risco a integridade física dos paciente e contribuindo para quedas, acidentes domésticos, lesões e agravamento do quadro de saúde que, por afetar em sua maioria os pacientes idosos, comumente evolui para o óbito, com 329 mil mortes registradas somente em 2019.

Neste contexto, o diagnóstico precoce contribui para a efetividade do tratamento,

melhorando a condução das intervenções do especialista quanto a evolução do estado clínico do paciente, visando uma melhora na sobrevida. Neste sentido, a pesquisa científica na área de extração de características de sinais, tanto de EEG quanto de voz, tem despertado crescente interesse, especialmente no contexto da classificação de pacientes com as patologias neurodegenerativas, tais como Doença de Alzheimer (PARIHAR; SWAMI, 2023), Doença de Parkinson (APPAKAYA *et al.*, 2023) mas também a Doença de Huntington (GALLEZOT *et al.*, 2022).

Pesquisas neste sentido são impulsionadas pela necessidade e urgência no desenvolvimento de abordagens diagnósticas cada vez mais precisas, não invasivas e acessíveis, para identificação precoce e tratamento destas doenças, permitindo intervenções terapêuticas mais eficazes. Outros desafios abrangem a análise e uso de sinais de EEG e de voz como biomarcadores, sendo o EEG uma valiosa ferramenta para investigar a atividade elétrica cerebral, enquanto a análise da voz oferece indícios sobre alterações neuromotoras e cognitivas, mesmo em diferentes fases de acometimento.

Por fim, destaca-se também a proposta de envolver os avanços tecnológicos no reconhecimento de padrão e métodos computacionais de aprendizado de máquina visando superar abordagens já apresentadas, afim de melhor contribuir para a extração e análise de características complexas presentes nesses sinais.

### **1.3 Trabalhos relacionados**

A literatura científica apresenta um corpo substancial de trabalhos relacionados, explorando diversas abordagens para a extração de atributos a partir de sinais biomédicos. Estudos têm se debruçado sobre a análise detalhada de características específicas, seja em sinais monocanais de voz por técnicas estatísticas ou acústicas avançadas, ou ainda no cenário multicanal com EEG, desvelando padrões complexos que têm como origem a atividade cerebral. Nesta seção, serão apresentados diversos trabalhos que oferecem contribuições significativas para o entendimento destes sinais e contribuem para o diagnóstico precoce de doenças neurodegenerativas.

O diagnóstico de doenças neurodegenerativas é um processo complexo que muitas vezes requer uma abordagem multifacetada para obter uma compreensão completa da condição do paciente. Tanto as abordagens monocanal quanto as de multicanal desempenham papéis importantes nesse processo. Este trabalho apresenta uma divisão em duas grandes seções, a Seção 1.3.1 apresenta abordagens de diagnóstico com foco em metodologias monocanais, por outro lado a Seção 1.3.2 apresenta avanços com sinais multicanais para classificação dos pacientes.

### 1.3.1 Metodologias em monocanal

No cenário de captação de sinais monocanal, os estudos sobre doenças neurodegenerativas investigam diferentes tipos de sinais biomédicos, tais como abordagens que utilizam acelerometria (MELLONE *et al.*, 2016; MURO-DE-LA-HERRAN *et al.*, 2014), que avalia informações sobre padrões de movimento e postura, útil na avaliação de distúrbios motores, há também cenários de captação de sinais de temperatura corporal, com abordagens utilizando termografia (SHAKARCHI, 2017), mas também sinais de voz, onde se destacam abordagens recentes, de baixo custo, cujo dispositivo de captação é amplamente acessível. Nesta seção, abordamos trabalhos relacionados a sinal monocanal, inspirados no estudo de sinais de voz.

#### 1.3.1.1 Abordagens relacionadas a Doença de Huntington

O estudo proposto por Perez *et al.* (2018) apresenta a classificação de indivíduos com o gene da HD usando sinais de voz. Os métodos envolvem a utilização de k-NN e DTW e a implementação de redes neurais profundas (DNN e LSTM). As descobertas indicam que as abordagens de aprendizagem profunda superam outras, alcançando uma acurácia média de 87%. As transcrições dos registros de áudios, geradas por humanos, forneceram resultados superiores para o teste de classificação, mas a qualidade da transcrição afetou a precisão. Isso ressalta a necessidade de métodos automatizados de extração de atributos da voz para esta análise.

Por outro lado, Riad *et al.* (2020) apresenta uma metodologia para investigação de indivíduos saudáveis e portadores do gene da doença de Huntington, e propõe um modelo de regressão linear para prever parâmetros das amostras que são de importância clínica para os diferentes indivíduos com a doença. Foram analisados pacientes com base nas características da fonação sustentada. Os resultados do trabalho quanto a classificação dos grupos se mostraram insatisfatórios, alcançando uma acurácia máxima de 56% e um F1-Score de 46% no melhor cenário. Além disso, o modelo apresentou desempenho pouco promissor com índice de correlação de 0,53, o que indica correlação fraca e pode indicar uma natureza potencialmente não linear dos dados, o que impõe maior dificuldade ao modelo quando em situações de classificação amostral.

O trabalho proposto por Romana *et al.* (2020) apresenta uma abordagem para análise da fala de indivíduos com a doença utilizando um conjunto de atributos denominado FVDM. Os autores destacam que a metodologia empregada pode aumentar significativamente a precisão da detecção da HD manifesta em estágio avançado, atingindo uma taxa de acurácia de 80%.

No entanto, o trabalho enfrenta desafios para distinguir HD nos estágios iniciais, indicando a necessidade de mais pesquisas para melhorar a precisão neste aspecto. Além disso, o tamanho da amostra de pacientes com HD em estágio avançado é relativamente pequeno, o que pode limitar a generalização dos resultados.

Por outro lado, o estudo de Vásquez-Correa *et al.* (2021), explorou o uso de modelos de aprendizado de máquina para classificar pacientes com doença de Huntington (DH) e doença de Parkinson (DP) com base em suas gravações de voz. Os resultados foram expressivos, demonstrando uma taxa de acurácia de 85,2% na classificação entre pacientes saudáveis em relação aos pacientes com HD. O estudo investigou a aprendizagem por transferência (conhecida também por *transfer learning*, do inglês) com CNNs para avaliar seu potencial para melhorar a detecção de doenças. O trabalho sugere uma forte correlação entre o estágio avançado da doença e o desempenho do modelo, com resultados menos eficazes nos estágios iniciais da HD.

Já Subramanian *et al.* (2023) desenvolveu uma metodologia para detectar a doença e avaliou em pacientes com estágios diferentes da HD. Os autores combinam três classificadores: Random Forest, Support Vector Machine e Logistic Regression. Os resultados indicaram uma acurácia de no máximo 73%. Este índice possivelmente decorre da dificuldade na melhor seleção aplicada nos atributos. Além disso, a classificação da doença apresentou resultados de baixa acurácia mesmo em casos de doença já manifestada, que deveria ser mais fácil de detectar, se comparado a trabalhos anteriores.

### **1.3.2 Metodologias multicanais**

Por outro lado, ao analisar características provenientes de múltiplos canais, tais como ocorre na atividade cerebral, os estudos oferecem uma abordagem abrangente, capturando nuances sutis que podem aprimorar a precisão diagnóstica e proporcionar uma compreensão mais abrangente das condições ao nível neurofisiológico. Uma revisão desses trabalhos aplicados a abordagem multicanal com eletroencefalograma proporciona uma perspectiva detalhada sobre os avanços e desafios no campo das abordagens multicanais para diagnóstico precoce.

#### **1.3.2.1 Abordagens relacionadas a Doença de Alzheimer**

A abordagem de Fiscon *et al.* (2018) combina o processamento de sinais de EEG com métodos supervisionados para classificar pacientes com Alzheimer, aplicando transformadas de Fourier e Wavelet para extrair características e em sua abordagem utiliza métodos clássicos

baseados em árvore de decisão, alcançando acurácia próxima a 83%. Este estudo, fomenta a análise no domínio da frequência, os autores apresentam uma análise utilizando ANOVA e Teste-T e reforçam que existem indicações mais evidentes nos coeficientes espectrais que contribuem para a melhor classificação das amostras, destacando a importância da análise das amostras no domínio da frequência.

Sob outra perspectiva, utilizando princípios da teoria dos grafos para a representação do sinal, Pineda *et al.* (2020) apresenta uma abordagem que constrói grafos para classificação de sinais de EEG. Em sua abordagem, o grafo é construído a partir do sinal no domínio do tempo, cuja matriz de incidência é normalizada e então convertida para uma matriz de transições de Markov. A partir desta estrutura, os autores extraem quatro atributos: coeficiente de agrupamento, modularidade de agrupamento, o comprimento médio de caminhos, a centralidade de vértices, e índice Estrada laplaciano. Cada atributo foi avaliado em sinais de pacientes em repouso, com olho aberto e fechado e alcança 100% de acurácia em 10% das amostras utilizadas para teste, 90% foram dedicadas para treino. Os autores relatam que o conjunto de pacientes é reduzido e carece de mais amostras para uma melhor avaliação quanto a abrangência do método.

Mais recentemente, Puri *et al.* (2023) apresentaram um conjunto de filtros, baseado em wavelet aplicado ao EEG para o diagnóstico automatizado do Alzheimer e do comprometimento cognitivo. Essa abordagem tenta explorar a relação entre o domínio do tempo e da frequência, tendo alcançado uma taxa de acerto superior a 90% para os conjuntos AD em relação aos saudáveis (HC, do inglês *Healthy Control*) e até 96% para a classificação AD em relação a deficiência cognitiva leve (MCI, do inglês *Mild Cognitive Impairment*). Resultados relevantes na classificação da patologia. Esse estudo realça a continuidade do uso de métodos baseados em wavelet para detecção de distúrbios neurodegenerativos.

A proposta de Puri *et al.* (2023), embora apresentando resultados expressivos, foi avaliada em um conjunto relativamente pequeno de pacientes e em todos os casos, apenas sinais de EEG registrados com os pacientes com os olhos fechados foram considerados, o que reduz as possibilidades de ruídos (artefatos) no sinal.

Em uma linha semelhante Miltiadous *et al.* (2023a) introduz uma arquitetura de rede convolucional para a detecção de Alzheimer em sinais de EEG. O estudo mostra que a combinação de redes convolucionais com camadas de transformadores de codificação, que é composta por um codificador e um decodificador, podem capturar efetivamente características complexas dos sinais de EEG. Em sua proposta, os sinais são filtrados em cinco faixas de

frequências específicas e destas são extraídas características de densidade espectral do sinal (PDS, do inglês *Power Spectral Density*) e conectividade de frequências baseada em coerência espectral. Os resultados embora promissores, 84% de acurácia para as classes AD vs. HC, são avaliados em apenas 13 pacientes com Alzheimer, carecendo de um teste mais abrangente com mais amostras de pacientes.

Por outro lado, Chen *et al.* (2023) propõem uma combinação de múltiplos atributos do sinal visando assim, promover uma melhora na separabilidades das classes Alzheimer em relação a pacientes saudáveis. O trabalho usa sinais de EEG em estado de repouso, combinando CNNs e Transformadores Visuais para superar as limitações de extração de características e melhorar a detecção de padrões anormais de sinais de EEG. Neste sentido os autores propõem a classificação das imagens, contendo os sinais de cinco diferentes bandas de frequência, construídos a partir de segmentos de 4 segundos dos canais de EEG. A metodologia apresenta taxas de acertos próximas a 87% para as classes AD vs HC. E próximas a 82% quando compara AD com a demência frontotemporal (DFT, do inglês *Frontotemporal dementia*).

Ainda avaliando técnicas de grafos com EEG, o estudo de Vicchietti *et al.* (2023) constrói grafos a partir das transições da cadeias de Markov e avalia com seis atributos extraídos de séries temporais (coerência wavelet, dimensão fractal, entropia quadrática, energia wavelet, e matriz de transição de Markov) nos sinais de EEG de 160 pacientes com Alzheimer e 24 saudáveis, em estado de repouso, em um momento com olho aberto e em outro com olho fechado. Os resultados indicam que tanto os sinais de EEG brutos quanto os filtrados por wavelet (bandas alfa, beta, teta e delta), analisados sob o domínio da frequência mostram que os métodos de análise de séries temporais testados, como a coerência wavelet e a matriz de Markov construída a partir do grafo, podem discriminar com resultados relevantes os pacientes com Alzheimer em relação a idosos saudáveis.

### 1.3.2.2 Abordagens relacionadas a Doença de Parkinson

O trabalho proposto por Anjum *et al.* (2020) introduziu o algoritmo LEAPD como uma nova abordagem destinada a diferenciar pacientes com a DP de grupos saudáveis. No conjunto de dados Iowa, o LEAPD demonstrou uma precisão promissora, alcançando 85,7% de acerto, acompanhada de uma Área sob a Curva Característica de Operação do Receptor (AUC) de 85,2%. As métricas adicionais incluem uma sensibilidade de 85,7% e uma especificidade de 85,7%, destacando uma boa consistência do LEAPD. Todavia, o estudo não apresenta

comparações com outros distúrbios neurodegenerativos que podem indicar limitações do LEAPD.

Já o trabalho proposto por Shah *et al.* (2020) sugere uma abordagem denominada DGHNet. Esta arquitetura de rede neural profunda, baseada nas propriedades dinâmicas e espectrais dos dados de EEG, alcança resultados notáveis, com taxas de acerto médias de 99,22%, especificidade de 99,46% e sensibilidade de 98,98%. A análise comparativa com modelos anteriores, como o EEGNet, destaca o desempenho superior do DGHNet, além de sua eficiência computacional, utilizando apenas 24% dos recursos em comparação com o EEGNet.

Ainda nessa perspectiva, Loh *et al.* (2021) propuseram um modelo de aprendizado profundo baseado na arquitetura 2D-CNN para a detecção automatizada da doença. A metodologia utiliza sinais de EEG para treinamento e teste do modelo, obtidos do conjunto de amostras "SanDiego". O estudo conseguiu diferenciar entre controles saudáveis de pacientes com Parkinson, alcançando uma taxa de acurácia de 99,46% em um dos experimentos realizados. Além disso, o estudo destacou a capacidade de detectar pacientes com DP, identificando se estavam ou não medicados. No entanto, o estudo indica um alto custo computacional para o processamento da CNN bidimensionais proposta, resultando em tempos de treinamento longos, além disso, a metodologia considerou apenas 15 indivíduos doentes.

Por outro lado, o trabalho apresentado por Aljalal *et al.* (2022) propõe uma metodologia baseada no Padrão Espacial Comum<sup>4</sup> para melhorar a detecção da doença através via sinais de EEG. O estudo utiliza dois conjuntos de dados, sendo o conjunto "SanDiego" um dos focos principais. O estudo divide a base em três diferentes combinações de pacientes, a metodologia alcança 95,76% para pacientes com PD sob medicação versus Saudável, já para o conjunto de pacientes com PD fora de medicação vs Saudável apresentou 99,41%.

O estudo proposto por Hamidi *et al.* (2023) aborda a detecção da doença de Parkinson por meio de sinais de EEG, propondo uma técnica chamada "*Forged Channels*" para representar os diferentes canais. A metodologia combina a transformação do sinal com uma rede neural profunda alcançando uma precisão de 90,32% e utiliza a metodologia de validação cruzada. Entretanto, os autores destacam que o modelo teve dificuldades de generalização pelo número reduzido de amostras utilizadas. Por outro lado, o artigo destaca que além das limitação impostas pelos métodos clínicos atuais, é possível o diagnóstico precoce automatizado.

Já o estudo proposto por Sugden e Diamandis (2023) apresenta uma abordagem de aprendizado profundo para a classificação inter-sujeito da Doença de Parkinson (DP) a partir de

---

<sup>4</sup> Técnica que visa melhorar a qualidade dos sinais ao reduzir a interferência de fontes não desejadas, especialmente em contextos onde múltiplos sensores são utilizados para aquisição de dados

dados de eletroencefalografia (EEG) em repouso. Utilizando a estratégia de validação cruzada, o modelo foi treinado e testado no conjuntos de dados Iowa. Os resultados demonstraram uma taxa de acerto robusta próximas a 82,80% e sugere uma boa eficácia do modelo. Apesar do bom desempenho na classificação, os autores destacam que são necessárias otimizações adicionais e validações em outros conjuntos de dados antes de considerar a implementação clínica definitiva do método.

Já abordagem proposta por Delfan *et al.* (2023), apresenta um modelo baseado em aprendizagem profunda para o diagnóstico da DP utilizando o sinal de eletroencefalograma (EEG) em estado de repouso. O objetivo do estudo é desenvolver um modelo automatizado que possa extrair características não lineares ocultas complexas do EEG e demonstrar a sua generalização em dados não vistos. O modelo é concebido utilizando um modelo híbrido, constituído por uma rede neural convolucional (CNN), uma unidade recorrente bidirecional<sup>5</sup> (Bi-GRU) para modelagem temporal, enfatizando partições relevantes do sinal EEG.

## 1.4 Objetivos

Como a abordagem proposta concentra-se na interseção de diferentes áreas, esse trabalho tem como objetivo geral estabelecer a eficácia do descritor de características de sinais biomédicos baseado em grafos direcionados sobre componentes espectrais. Além de propor o método de extração, indica também seus resultados com aplicação em diferentes patologias com foco em doenças neurodegenerativas, e, conseqüentemente, propõe a flexibilidade da metodologia proposta em cenários de aplicação real em mono e multicanal, com vistas no potencial de associação do método com algoritmos de classificação baseados em tensores.

### 1.4.1 *Objetivos Específicos*

- Propor uma estrutura de análise do sinal que envolva operações simples e baseada em grafo com vistas a redução de custo com operações matemáticas mais onerosas, além de estabelecer uma estrutura compatível com operações em grafo;
- Propor uma abordagem tensorial, capaz de representar as amostras em tensores tridimensionais assim ser utilizada em conjunto com redes neurais profundas, para que além do potencial quanto a classificação, seja compatível com técnicas recentes de transferência de

<sup>5</sup> Arquitetura de unidade recorrente utilizada em redes neurais, especialmente em tarefas de processamento sequencial, como processamento de linguagem natural e séries temporais.

aprendizado;

- Propor uma metodologia de extração que seja capaz de processar sinais tanto monocanal quanto multicanal;
- Estabelecer um descritor baseado em grafo com potencial de discriminação para cenário de classificação binário e multi-classe;
- Demonstrar a flexibilidade da metodologia proposta em cenários de aplicação real com amostras de doenças neurodegenerativas, com vistas ao potencial de associação da extração com algoritmos de classificação baseados em tensores.

## 1.5 Produção científica

Durante este projeto de doutorado, os seguintes trabalhos científicos foram produzidos, a saber:

- Trabalhos diretamente relacionados a esta pesquisa:
  - **Medeiros, A. G.**, et al. "Bi-Dimensional Approach Based on Graph Neural Network for Alcoholism Predisposition Detection via EEG signals". International Joint Conference on Neural Networks (IJCNN). IEEE, 2023. (Qualis A1)
  - Zhang, H., Silva, F. H., Ohata, E. F., **Medeiros, A. G.**, Rebouças Filho, P. P. "Bi-dimensional approach based on transfer learning for alcoholism pre-disposition classification via EEG signals". Frontiers in Human Neuroscience, 14, 365, 2020.(Qualis A2)
  - **Medeiros, A. G.**, et al. "New Feature Extractor based on Directed Graphs in Spectral Components for Detecting Alzheimer's Disease from multi-channel EEG signal". Pattern Recognition Letters. Elsevier. (Em revisão)
  - **Medeiros, A. G.**, et al. "New Feature Extractor based on Directed Graphs in Spectral Components for Detecting Huntington's Disease via Audio Signals". Measurement. Elsevier. (Em revisão)
- Outras publicações em periódicos internacionais:
  - de Souza Rebouças, E., de Medeiros, F. N. S., Marques, R. C. P., Chagas, J. V. S., Guimaraes, M. T., Santos, L. O., **Medeiros, A. G.**, Peixoto, S. A. "Level set approach based on Parzen Window and floor of log for edge computing object segmentation in digital images". Applied Soft Computing, 105, 107273, 2021. (Qualis A1)
  - Peixoto, S. A., Vasconcelos, F. F., Guimarães, M. T., **Medeiros, A. G.**, Rego, P.

- A., Neto, A. V. L., Reboucas Filho, P. P. "A high-efficiency energy and storage approach for IoT applications of facial recognition". *Image and Vision Computing*, 96, 103899, 2020. (Qualis A2)
- Peixoto, S. A., **Medeiros, A. G.**, Hassan, M. M., Dewan, M. A. A., Albuquerque, V. H. C. D., Filho, P. P. R. "Floor of log: a novel intelligent algorithm for 3D lung segmentation in computer tomography images". *Multimedia Systems*, 1-13, 2020. (Qualis A2)
  - Peixoto, Solon A. et al. "A high-efficiency energy and storage approach for IoT applications of facial recognition". *Image and Vision Computing*, v. 96, p. 103899, 2020. (Qualis A2)
  - Vasconcelos, F. F. X., **Medeiros, A. G.**, Peixoto, S. A., Reboucas Filho, P. P. "Automatic skin lesions segmentation based on a new morphological approach via geodesic active contour". *Cognitive Systems Research*, 55, 44-59, 2019. (Qualis A1)
  - **Medeiros, A. G.**, et al. "A new fast morphological geodesic active contour method for lung CT image segmentation." *Measurement* 148, 2019. (Qualis A1)
- Outras publicações em conferências internacionais:
    - **Medeiros, A. G.**, et al. "Fast Stroke Lesions Segmentation Based on Parzen Estimation and Non-uniform Bit Allocation in Skull CT Images". *International Conference on Intelligent Systems Design and Applications*. Cham: Springer Nature Switzerland, 2022. (Qualis B1)
    - Santos, Lucas de O., **Medeiros, A. G.**, and Pedro P. Rebouças Filho. "Fast fully automatic stroke lesions segmentation based on Parzen estimation and  $\mu$ -law in skull CT images". *Congresso Brasileiro de Automática-CBA*. Vol. 3. No. 1. 2022. (Qualis B4)
    - Nunes, Virgínia Xavier, et al. "A novel web platform for covid-19 diagnosis using x-ray exams and deep learning techniques." *2021 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2021. (Qualis A1)
    - de Souza, João W. Mendes, et al. "Fingerprint Classification Based on the Henry System via ResNet". *International Conference on Systems, Signals and Image Processing*. Cham: Springer International Publishing, 2021. (Qualis A1)
    - **Medeiros, A. G.**, et al. "An approach to BI-RADS uncertainty levels classification

via deep learning with transfer learning technique". 2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS). IEEE, 2020. (Qualis B1)

## **1.6 Organização da Tese**

Este trabalho apresenta na Seção 1.3 os trabalhos recentes relacionados ao estudo de sinais para a classificação de pacientes acometido com doenças neurodegenerativas, com ênfase em três patologias dentre as mais incidentes, abrangendo abordagens que utilizam sinais mono e multicanais. O Capítulo 2 traz consigo os principais conceitos utilizados nesse trabalho, construindo um arcabouço teórico que deu base para este estudo. O Capítulo 3 apresenta a metodologia proposta, incluindo a explanação de como o descritor processa o sinal, contemplando também a proposta dos marcadores espectrais também desenvolvidos nesta tese. Os resultados em conjunto de dados sintéticos são apresentados e discutidos no Capítulo 4. Por conseguinte, no Capítulo 5 é proposta a aplicação do método com sinais de voz registrados por pacientes reais para a classificação da Doença de Huntington; já nos Capítulos 6 e 7, são propostas duas aplicações em cenário com amostras reais, em duas aplicações distintas que usam sinais de EEG para classificação de pacientes com a Doença de Alzheimer e Parkinson, respectivamente.

## 2 FUNDAMENTAÇÃO TEÓRICA

A metodologia de extração de atributos via sinais biológicos que é apresentada neste trabalho encontra sua inspiração matemática em diferentes avanços apresentados na literatura. Seja a partir do ponto de vista da abstração do problema em questão e sua representação em estruturas matemáticas como grafos, seja na construção de tensores espectrais que, sob uma construção fundamentada na análise espectral tradicional, poderão ser usados como insumo para treinamento de modelos de classificação com base em tecnologias recentes, como é o caso de modelos de classificação baseados em aprendizado profundo. Neste contexto, a seção que segue apresenta os fundamentos teóricos que balizarão o desenvolvimento da metodologia, bem como diferentes técnicas que também serão utilizadas como mecanismos de avaliação a fim de averiguar o desempenho da abordagem proposta.

### 2.1 Teoria dos Grafos

A teoria dos grafos é um ramo da matemática que estuda a relação entre vértices e arestas, como uma ferramenta para mapear situações do mundo real de modo matemático e com uma estrutura representativa flexível. O conceito fundamental de grafos teve suas origens no século XVIII com o trabalho de Leonhard Euler. Em 1736, Euler resolveu o famoso problema das Pontes de Königsberg, onde ele representou a cidade e suas pontes como um grafo, dando início à teoria dos grafos (EULER, 1741; DIESTEL; KUHN, 2003; WEST *et al.*, 2001).

Como descreve Diestel e Kuhn (2003), o problema em questão envolvia a cidade de Königsberg, que era atravessada pelo rio Pregel, dividindo a cidade em quatro áreas de terra ligadas por sete pontes. O desafio era determinar se era possível passear por toda a cidade atravessando cada ponte uma única vez e retornar ao ponto de partida. Para descobrir isso, pode-se representar as terras por pontos e as pontes por linhas; então, as condições dadas seriam satisfeitas exatamente quando o desenho resultante seria tal que se pudesse caminhar de um ponto a outro sem tirar a caneta do papel (DIESTEL; KUHN, 2003). Assim, Euler então estabelecia as bases para a teoria dos grafos.

#### 2.1.1 Grafos

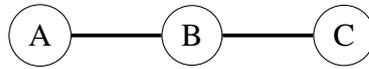
Um grafo  $G$  é uma estrutura matemática que consiste em um conjunto de vértices  $V$  e um conjunto de arestas  $E$ . Formalmente, um grafo é definido como  $G = (V, E)$ , onde  $V$  é

um conjunto não vazio de vértices e  $E$  é um conjunto de pares de vértices, representando as arestas (WEST *et al.*, 2001; DIESTEL; DIESTEL, 2017). Seja  $V = \{v_1, v_2, \dots, v_k\}$  um conjunto de vértices, vale a seguinte definição:

$$E \subseteq \{\{v_i, v_j\} \mid v_i, v_j \in V\} \quad (2.1)$$

onde  $\{\{v_i, v_j\} \mid v_i, v_j \in V\}$  denota o conjunto de todos os pares não ordenados de elementos distintos de  $V$ , como descreve West *et al.* (2001).

Por exemplo, se considerarmos  $V = \{A, B, C\}$  em que os vértice de  $V$  estão ligados por uma aresta do conjunto  $E$ , considerando  $E$  definido como  $E = \{(A, B), (B, C)\}$ , temos então o grafo  $G = (V, E)$ , definido graficamente por:



Nota-se que neste caso não há uma aresta que ligue diretamente os vértices  $A$  e  $C$ , visto que não existe tal definição como par de vértices no conjunto  $E$ . Vale destacar que o conjunto  $E$ , neste caso, é definido por pares de vértices onde a ordem não importa, isto é, deseja-se apenas indicar que há uma ligação entre vértices independentemente do sentido.

### 2.1.2 Digrafos

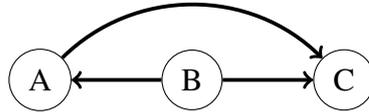
Por outro lado, como apresenta Diestel e Diestel (2017), um digrafo é um grafo direcionado, isto é, um grafo em que cada aresta tem um sentido associado. Formalmente, um grafo direcionado é representado por  $G = (V, E)$ , onde  $V$  é o conjunto de vértices e  $E$  é o conjunto de pares ordenados de vértices, neste sentido, a ordem dos pares é importante para a definição das relações entre vértices. Por exemplo, considere um grafo representando as ruas de uma cidade como arestas e os cruzamentos como vértices. Nem todas as ruas são de mão dupla. Ao se estudar rotas de ônibus, por exemplo, é necessário considerar se as ruas são de mão única, isto é, permitem fluxo apenas no sentido  $(v_i \rightarrow v_j)$ , ou se são de mão dupla  $(v_i \leftrightarrow v_j)$ , para o primeiro caso, a ordem do par importa.

Neste sentido, como detalha Bang-Jensen e Gutin (2008), seja  $G$  um grafo direcionado, considerando  $V = \{v_1, v_2, \dots, v_k\}$  um conjunto de vértices, vale a seguinte definição:

$$E \subseteq \{\{v_i, v_j\} \mid v_i, v_j \in V \text{ e } \{v_i, v_j\} \neq \{v_j, v_i\}\} \quad (2.2)$$

onde  $\{v_i, v_j\} \neq \{v_j, v_i\}$  denota a diferença de sentido entre as arestas.

Por exemplo, se considerarmos  $V = \{A, B, C\}$  em que os vértice de  $V$  estão ligados por arestas direcionadas do conjunto  $E$ , considerando  $E$  definido como  $E = \{(B, A), (B, C), (A, C)\}$ , temos então o grafo  $G = (V, E)$ , definido graficamente por:



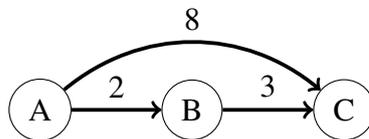
Note que no exemplo as arestas seguem o mesmo sentido da ordem do par de vértices dentro do conjunto  $E$ . Destaca-se também que não há um relacionamento partindo do vértice C para os demais vértices, o sentido das setas demonstra esta relação.

Já um digrafo de arestas ponderadas é um grafo direcionado em que cada aresta possui um peso associado. Formalmente, um grafo direcionado de arestas ponderadas é representado por  $G = (V, E, W)$ , onde  $V$  é o conjunto de vértices,  $E$  é o conjunto de pares ordenados de vértices representando as arestas direcionadas, e  $W$  é uma função de peso que atribui um peso a cada aresta de  $V$ . O conjunto de arestas  $E$ , portanto, pode ser definido pela relação:

$$E \subseteq \{(v_i, v_j) \mid v_i, v_j \in V\}, \quad W : E \rightarrow \mathbb{R} \quad (2.3)$$

onde  $\mathbb{R}$  representa os números reais e a relação define para cada par em  $E$  um valor real associado que indica o peso da aresta.

Por exemplo, se considerarmos  $V = \{A, B, C\}$  em que os vértice de  $V$  estão ligados por arestas direcionadas do conjunto  $E$  tal que  $E = \{(A, B, 2), (B, C, 3), (A, C, 8)\}$ , temos então o grafo  $G = (V, E)$  definido graficamente por:

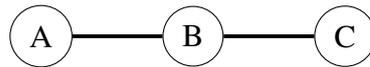


Existem várias maneiras de representar grafos, cada uma com suas vantagens em diferentes contextos. As representações mais comuns incluem: Matriz de Adjacência, Matriz Laplaciana, Matriz Laplaciana Normalizada, Matriz Laplaciana Combinatorial Direcionada. Cada uma destas estruturas descritas nas seções a seguir.

### 2.1.3 Matriz de Adjacência

No sentido de encontrar mecanismos para representar o grafo está a Matriz de Adjacência. Diestel e Diestel (2017) definem esta matriz como uma forma de representar grafos usando uma matriz  $A$ , onde  $A[i][j]$  é 1 se existe uma aresta entre os vértices  $i$  e  $j$ , e 0 caso contrário. Para um grafo não direcionado, a matriz é simétrica.

Como exemplo, considere o grafo  $G_1=(V,E)$ , com três vértices:



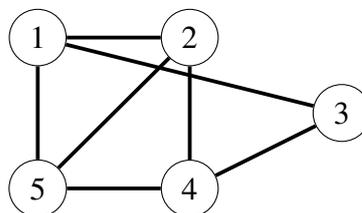
Neste caso, a Matriz de Adjacência  $A$  do grafo  $G_1$  é dada por:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (2.4)$$

Assim, como descreve West *et al.* (2001), a matriz de adjacência  $A$  reflete as conexões diretas entre os vértices do grafo. Além disso, a matriz é simétrica em relação à sua diagonal principal. Isso reflete a natureza bidirecional das arestas em grafos não direcionados, também facilita o acesso rápido à informação de adjacência, tornando-a eficiente para verificar a existência de uma aresta entre dois vértices. A matriz também apresenta propriedades que podem ser exploradas como representação de problemas na álgebra linear uma vez que os autovalores e autovetores dessa matriz podem ser usados em estudos espectrais de grafos.

### 2.1.4 Matriz Laplaciana

Já a matriz laplaciana, de acordo com Diestel e Diestel (2017), é útil para estudar propriedades algébricas de grafos. Se  $D$  é a matriz diagonal dos graus dos vértices e  $A$  é a matriz de adjacência, então a matriz laplaciana  $L$  é dada por  $L = D - A$ . Como exemplo, considere o grafo  $G_2=(V,E)$ , com vértices não direcionados e não ponderados:



Neste caso, com a matriz diagonal  $D$  dos graus, temos então a partir da subtração da Matriz  $D$  em relação a matriz de adjacência, a Matriz de Laplaciana  $L$  do grafo  $G_2$ , ambas descritas como:

$$D = \begin{bmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix} \quad (2.5)$$

$$L = \begin{bmatrix} 3 & -1 & -1 & 0 & 0 \\ -1 & 3 & 0 & -1 & -1 \\ -1 & 0 & 2 & -1 & 0 \\ 0 & -1 & -1 & 3 & -1 \\ 0 & -1 & 0 & -1 & 2 \end{bmatrix} \quad (2.6)$$

De acordo com Dong *et al.* (2016), a matriz laplaciana reflete as propriedades locais do grafo. Os valores diagonais representam o grau dos vértices, e os valores fora da diagonal indicam a conectividade entre os vértices. Esta matriz também carrega consigo informações para analisar propriedades espectrais e de conectividade do grafo. A diagonal principal da matriz, por exemplo, contém os graus dos vértices que facilita a identificação de alta ou baixa conectividade do vértice. Além disso, o segundo menor autovalor é conhecido como o autovalor de *Fiedler*<sup>1</sup> que representa características quanto a conectividade do grafo.

### 2.1.5 Matriz Laplaciana Combinatória

Já a matriz Laplaciana Combinatória de um grafo é uma representação matricial que fornece informações cruciais sobre sua estrutura, como descreve West *et al.* (2001). Em termos

<sup>1</sup> O autovalor de Fiedler ( $\Lambda_2$ ) fornece informações sobre a conectividade do grafo. Se o grafo é conexo  $\Lambda_2 > 0$ , já que  $\Lambda_2 = 0$  o grafo possui mais de uma componente. A forma dos autovetores associados  $\Lambda_2$  pode ser usada para particionar o grafo em dois conjuntos aproximadamente igualmente conectados.

matemáticos, os elementos da matriz Laplaciana Combinatória são dados por:

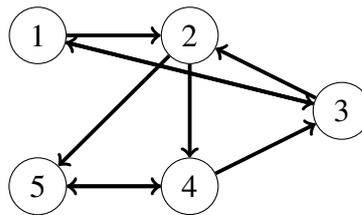
$$L[i][j] = \begin{cases} d_i & \text{se } i = j \text{ (diagonal principal)} \\ -1 & \text{se } i \neq j \text{ e } i \text{ e } j \text{ são adjacentes} \\ 0 & \text{se } i \neq j \text{ e } i \text{ e } j \text{ não são adjacentes} \end{cases} \quad (2.7)$$

onde  $d_i$  é o grau do vértice  $i$ .

A matriz Laplaciana Combinatória abrange características importantes do grafo. West *et al.* (2001) ressalta que há características relevantes, tal como a simetria e a bilateralidade das relações entre os vértices, enquanto a soma das linhas e colunas destaca a distribuição balanceada dos graus dos vértices. Os autovalores não-negativos e o nulo próprio indicam propriedades estruturais fundamentais. Assim, a matriz Laplaciana Combinatória é uma ferramenta poderosa para análise espectral e compreensão do grafo.

### 2.1.6 Matriz Laplaciana Combinatorial Direcionada

Por outro lado, a matriz laplaciana combinatorial direcionada (LCD) descrita por West *et al.* (2001), é obtida pela fórmula  $LCD = D^{in} - A$ , onde  $D^{in}$  é a matriz diagonal dos graus de entrada dos vértices e  $A$  é a matriz de adjacência. Considere o grafo  $G_4 = (V, E)$ , tal que  $V = \{A, B, C, D, E\}$ , com arestas direcionadas:



em que a matriz diagonal ( $D^{in}$ ) e a matriz de adjacência direcionada ( $A$ ), bem como a Matriz Laplaciana Combinatorial Direcionada (LD) são definidas como:

$$D^{in} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix} \quad (2.8)$$

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad (2.9)$$

$$LD = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 2 & 0 & -1 & -1 \\ -1 & -1 & 2 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 2 \end{bmatrix} \quad (2.10)$$

Como apresenta Chung (2005), a matriz laplaciana combinatorial direcionada reflete as propriedades direcionadas do grafo, levando em consideração apenas os graus de entrada dos vértices. Os termos não nulos fora da diagonal da matriz LCD indicam as relações de conectividade entre os nós, tais termos também podem indicar a presença e mensurar os ciclos no grafo. Além disso a matriz possui informações quanto ao segundo menor autovalor que indica o número de componentes fracamente conectados no grafo.

## 2.2 Transformada Discreta de Fourier (DFT)

A Transformada Discreta de Fourier (DFT), descrita como algoritmo computacional em formato discreto por Cooley e Tukey (1965), é uma transformação matemática que converte um sinal discreto no domínio do tempo em um sinal discreto no domínio da frequência. Isso significa que ela nos permite analisar um sinal e determinar quais componentes de frequência estão presentes bem como suas respectivas magnitudes.

### 2.2.1 Definição Matemática da DFT:

Neste sentido, do ponto de vista matemático, Press (2007) define a DFT de um sinal discreto  $x[n]$ , onde  $n$  representa o índice de amostragem, como a seguinte expressão matemática:

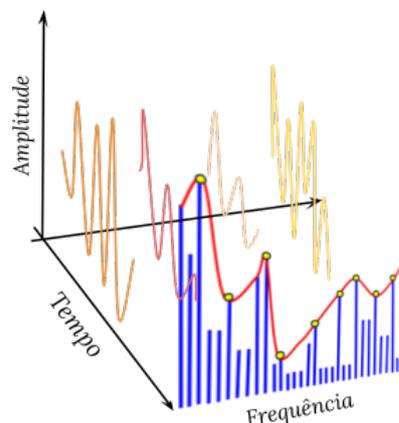
$$X[k] = \sum_{n=0}^{N-1} x[n] \cdot e^{-j\frac{2\pi}{N}kn}, \quad k = 0, 1, 2, \dots, N-1 \quad (2.11)$$

onde  $X[k]$  é a componente de frequência na posição  $k$ ,  $x[n]$  é o valor do sinal no tempo  $n$ ,  $N$  é o número total de amostras no sinal e  $j$  é a unidade imaginária ( $i$ ). A Figura 1 apresenta uma ilustração comparativa do sinal no domínio do tempo, sinal em vermelho, em relação a sua representação no domínio da frequência, sinal em azul, com suas respectivas frequências decompostas, sinais em outras cores.

A DFT é uma ferramenta matemática fundamental na análise de sinais discretos. Suas propriedades são essenciais para compreender como ela opera em diferentes contextos. Esta transformada apresenta propriedades relevantes em aplicações com sinais de uma dimensão, especialmente na análise de sinais biológicos e na modelagem de sistemas biológicos, como apresenta Raugei *et al.* (2006). Dentre suas principais propriedades, destacam-se:

- Linearidade, já que a DFT é uma transformação linear;
- Periodicidade, a DFT é periódica em seu domínio da frequência, com período igual ao número de amostras no sinal;
- Convolução no Domínio do Tempo, em que a multiplicação de dois sinais no domínio da frequência corresponde à convolução dos sinais no domínio do tempo;
- Simetria Hermitiana, a DFT de um sinal é simétrica em relação ao seu centro (em termos de frequências positivas e negativas)

Figura 1 – Representação do sinal sob duas perspectivas: domínio do tempo (em laranja), com suas respectivas frequências decompostas (em outras cores); e no domínio da frequência, componentes espectrais representados por barras (em azul)



Fonte: Elaborada pelo autor.

### 2.2.2 *Relação da DFT com Ruídos nos Sinais*

A Transformada Discreta de Fourier desempenha um papel essencial na análise de sinais, sendo especialmente valiosa na presença de ruídos (OPPENHEIM, 1999). Explorar aspectos da DFT no sentido de melhor lidar com ruídos em sinais, com foco em exemplos práticos, notadamente em sinais biológicos, contribui para o desenvolvimento de tecnologias no apoio ao diagnóstico de patologias cerebrais, principalmente com sinais de Eletroencefalograma (MOOTOO *et al.*, 2023; NAJEEB *et al.*, 2021; MATHUR; CHAKKA, 2020).

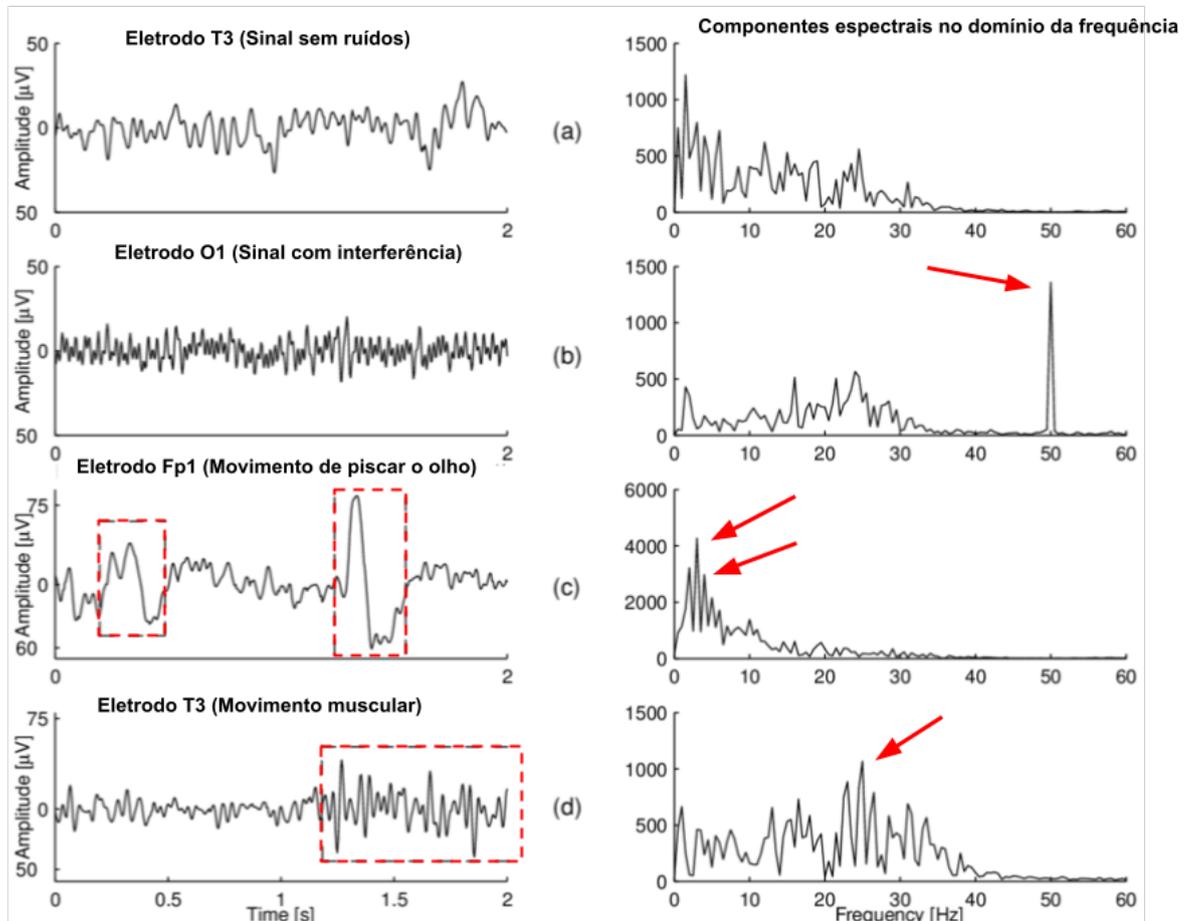
Neste sentido, a decomposição de um sinal em suas componentes de frequência é particularmente útil ao lidar com sinais biológicos, como é o caso do Eletroencefalograma, em que muitas vezes a coleta dos sinais no paciente está sujeita a interferências eletromagnéticas, ou de atividades biológicas do próprio paciente como piscar ou abrir o olho, além de outros tipos de ruídos envolvendo o equipamento e eletrodos.

Ao aplicar a DFT a sinais biológicos, é possível identificar componentes espectrais associadas tanto ao sinal de interesse quanto aos ruídos indesejados. Considere um sinal EEG, por exemplo, a DFT pode revelar picos em determinadas frequências que correspondem às atividades cerebrais desejadas, enquanto componentes fora dessas frequências podem ser identificadas como ruídos.

A Figura 2, por exemplo, apresenta diferentes sinais de EEG com suas respectivas representações no domínio da frequência, à direita. Em todos os casos temos exemplos de sinais em diferentes eletrodos a partir de um exame de EEG. Na primeira linha, um exemplo de sinal sem ruídos e sua versão no domínio da frequência. Já na segunda linha, um sinal obtido pelo eletrodo O1 com interferência da linha de energia, já na figura da terceira linha dois ruídos também conhecidos como artefatos, produzidos por movimentos naturais do paciente, tais como movimento ocular e piscar o olho, por fim, no último caso uma contração muscular do paciente causa ruídos e em todos os casos tais ruídos puderam ser melhor representados como componente espectral no domínio da frequência.

Diante disso, é possível dizer que a estimação espectral de ruído, através da DFT, é crucial em aplicações onde é necessário compreender as características do sinal de interesse em relação ao ruído presente nos sinais biológicos. Isso contribui para o desenvolvimento de estratégias mais robustas ou mais eficazes na remoção ou redução de ruído. Deste modo, ao compreender as componentes espectrais e aplicar estratégias adequadas de filtragem e análise do sinal, é possível extrair informações significativas desses sinais, que são essenciais em aplicações

Figura 2 – Sinais EEG perturbados por ruído e artefatos e sua potência espectral correspondente. (a) sinal EEG sem ruídos registrado no eletrodo T3. (b) sinal EEG, registrado no eletrodo O1, perturbado pelo ruído da linha de energia. (c) Sinal registrado no eletrodo Fp1 contendo movimento ocular (esquerda) e piscar de olhos artefatos (à direita). (d) Sinal registrado no eletrodo T3, contendo artefato de movimento muscular.



Fonte: Adaptada pelo autor a partir de Molina (2004).

clínicas e de pesquisa

### 2.3 Extração clássica de atributos em sinal unidimensional

Diferentes extratores de atributos são apresentados na literatura para análise de sinais unidimensionais, dentre os quais se destacam os Coeficientes Mel-Cepstrais (MFCCs) (LOGAN, 2000), além da magnitude dos coeficientes da Transformada Discreta de Fourier (Magnitude dos coeficientes). Outros extratores incluem atributos estatísticos que consideram medidas: média, variância, desvio padrão, assimetria e curtose, além do extrator estatístico de ordem superior (OHS) e aqueles baseados em *wavelets*. Esta seção, apresenta portanto uma descrição de cada

extrator clássico, suas principais características e formulação matemática.

### 2.3.1 Extratores baseados em estatísticas do sinal:

Considerando o domínio da frequência, Logan (2000) descreve os coeficientes Mel-Cepstrais, também conhecidos como coeficientes cepstrais de Mel (MFCCs, do inglês *Mel-Frequency Cepstral Coefficients*), são características amplamente utilizadas em processamento de sinais de áudio, processamento de fala e reconhecimento de fala. Esses coeficientes foram originalmente desenvolvidos para modelar a percepção auditiva humana e são especialmente eficazes na representação das características espectrais de um sinal de áudio.

Ainda no domínio da frequência, originalmente Cooley e Tukey (1965) descreve a magnitude da DFT como uma transformação matemática que converte um sinal discreto, composto por amostras de amplitude finita, do domínio do tempo para o domínio da frequência. Em outras palavras, ela analisa um sinal no domínio das frequências, decompondo-o em suas componentes senoidais. A magnitude da DFT é uma medida da contribuição de diferentes frequências em um sinal, e pode ser definida como:

$$X(k) = \sum_{n=0}^{N-1} x(n)e^{-j\frac{2\pi}{N}kn} \quad (2.12)$$

onde  $x(n)$  é a amostra do sinal no instante  $n$ ,  $N$  é o número total de amostras, e  $X(k)$  é o coeficiente na frequência  $k$ .

Além do processamento de características espectrais, existem também os extratores estatísticos, no contexto de processamento de sinais e análise de dados, Hastie *et al.* (2009) descreve que atributos estatísticos calculam diversas medidas para caracterizar propriedades importantes de um conjunto de dados. Essas estatísticas são utilizadas para descrever a distribuição, tendência central, dispersão e outras características relevantes dos dados.

Em diferentes aplicações, Hastie *et al.* (2009) destaca que os extratores estatísticos podem ser empregados para analisar e representar propriedades específicas de sinais ou conjuntos de dados. Os atributos estatístico consideram o cálculo de várias estatísticas descritivas de um conjunto de dados. Considere, o conjunto  $X$  de dados contendo  $N$  amostras,  $X = \{x_1, x_2, \dots, x_N\}$ , é possível elencar os seguintes atributos estatísticos que incluem: Média ( $\mu$ ), Variância ( $\sigma^2$ ), Desvio Padrão ( $\sigma$ ), Assimetria ( $\gamma$ ) e Curtose ( $\kappa$ ). Todos estes são descrições estatísticas clássicas de um determinado conjunto.

Ainda dentro do contexto de medidas estatísticas, um dos pioneiros em propor análise

de medidas estatísticas foi Nandi (1999). Em seus estudos, destaca que o extrator estatístico de Ordem Superior (OHS) é uma ferramenta analítica que vai além das estatísticas tradicionais de primeiro e segundo momentos, tais como média e variância, ao incluir momentos de ordem mais elevada na descrição da distribuição de dados. Esses momentos superiores proporcionam informações adicionais sobre a forma e o padrão da distribuição, sendo particularmente úteis em contextos onde propriedades mais complexas dos dados precisam ser identificadas e avaliadas.

### 2.3.2 Extratores baseados em wavelets

As *wavelets* são funções matemáticas essenciais em análise de sinais e processamento de imagens, permitindo a representação e análise de características em diferentes escalas. Rioul e Vetterli (1991) detalham que dada uma função mãe (*wavelet* mãe)  $\psi(t)$ , a decomposição de um sinal contínuo  $f(t)$  em coeficientes de alta e baixa frequência é dada por:

$$Wf(a, b) = \int_{-\infty}^{\infty} f(t) \cdot \psi_{a,b}(t) dt, \quad (2.13)$$

onde  $a$  e  $b$  são parâmetros de escala e deslocamento, respectivamente, e  $\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right)$  é a *wavelet* dilatada e deslocada.

Para sinais discretos, a formulação da transformada discreta de *wavelet* (DWT) toma como base uma sequência  $x[n]$ , a DWT é então definida como:

$$Wf(a, b) = \sum_n x[n] \cdot \psi_{a,b}[n], \quad (2.14)$$

onde  $\psi_{a,b}[n] = \frac{1}{\sqrt{a}} \psi\left(\frac{n-b}{a}\right)$  é a *wavelet* dilatada e deslocada para sinais discretos.

As famílias específicas de *wavelets*, como Coiflets (COIF), Symlets (SYM), Daubechies (DB), e Haar, são escolhidas com base em suas propriedades distintas, que desempenham papéis cruciais em diversas aplicações científicas (RIOUL; VETTERLI, 1991). Dentre as principais famílias de *wavelets*, destacam-se:

- Haar: *wavelet* Haar é simples e bastante eficaz na detecção de descontinuidades nos sinais.
- DB: são conhecidas por proporcionar uma boa compactação de energia, o que significa que um número relativamente pequeno de coeficientes é suficiente para representar eficientemente o sinal.
- SYM: rojetadas para terem uma simetria aprimorada em comparação com as Daubechies, tornando-as mais adequadas para sinais que podem não ser completamente simétricos.
- COIF: projetadas para compactar a energia do sinal de forma eficiente, além de aproximar sinais contínuos suaves.

Diversos trabalhos fazem uso dos extratores baseados em *wavelets*, seja para extração de atributos como são os trabalhos de Crovato e Schuck (2007), Fong *et al.* (2012) ou para a filtragem de ruídos em Liu *et al.* (2023), para a classificação de voz ou patologias por sinais de voz. Por outro lado, ainda no contexto de voz, o extrator MFCC, como apresenta Sol *et al.* (2023), além da Magnitude dos coeficientes de Fourier também são empregados neste mesmo contexto (APPAKAYA *et al.*, 2023), o que indica mais uma vez a necessidade da investigação minuciosa quanto a adequação do extrator em cada experimento e aplicação.

Quanto ao tamanho da representação de cada extrator, para fins de comparação, a Tabela 1 apresenta o número de atributos obtido por cada extrator, considerando a base de sinais de voz para a classificação da doença de Huntington, que será descrito no Capítulo 5.

Tabela 1 – Extratores de atributos clássicos e número total de atributos com base nos experimentos sobre amostras de voz de pacientes saudáveis e com a doença de Huntington.

Extrator de atributos	Número de atributos
Extrator de Coeficientes Mel-Cepstrais	1049600
Magnitude dos Coeficientes de Fourier	1049600
Atributos Estatísticos	5
OHS	4
Wavelet COIF1	262404
Wavelet COIF11	262456
Wavelet COIF13	262467
Wavelet COIF15	262477
Wavelet COIF3	262414
Wavelet COIF5	262425
Wavelet COIF7	262435
Wavelet COIF9	262446
Wavelet DB1	262400
Wavelet DB10	262416
Wavelet DB15	262425
Wavelet DB20	262434
Wavelet DB25	262442
Wavelet DB30	262451
Wavelet DB35	262460
Wavelet DB5	262407
Wavelet Haar	262400
Wavelet SYM10	262416
Wavelet SYM12	262420
Wavelet SYM14	262423
Wavelet SYM16	262427
Wavelet SYM2	262402
Wavelet SYM4	262406
Wavelet SYM6	262409
Wavelet SYM8	262413

Fonte: Elaborada pelo autor.

### 2.3.3 Extratores específicos para sinais de voz

Dentro do contexto do tratamento e análise de sinais de voz está a ferramenta conhecida como OpenSMILE (do inglês, *Open-Source Media Interpretation by Large feature-space Extraction*), desenvolvido pela Universidade de Passau, é uma biblioteca de código aberto que se destaca na extração de características de sinais de áudio, com foco especial em atributos vocais (EYBEN *et al.*, 2010). Esta ferramenta apresentou resultados promissores em trabalhos anteriores para extração de atributos e classificação de patologias a partir da análise de sinais de voz (GUIMARÃES *et al.*, 2020; ALMEIDA *et al.*, 2019; VAICIUKYNAS *et al.*, 2017).

A abordagem modular do OpenSMILE, de acordo Eyben *et al.* (2010), oferece um conjunto diversificado de ferramentas para análise de sinais de áudio em larga escala. Para fins de comparação, este estudo considera diferentes conjuntos de atributos extraídos dos sinais de voz gerados pela ferramenta. Neste sentido, elencamos doze diferentes configurações de conjuntos de atributos, que correspondem a composição de diferentes extratores e que foram utilizados em desafios dentro do contexto de análise de sinais de voz (SCHULLER *et al.*, 2011; SCHULLER *et al.*, 2013; SCHULLER *et al.*, ).

Os extratores *avec2011* e *avec2013* foram desenvolvidos para a tarefa de extração de características vocais no âmbito da competição INTERSPEECH 2011 (SCHULLER *et al.*, 2011). Este conjunto incorpora extratores tais como MFCCs, função de autocorrelação do sinal, além de propriedades estatísticas temporais como média, variância e curtose do sinal. A versão de 2013 apresenta-se como uma evolução do *avec2011*, incluindo outros atributos extraídos do sinal, tais como variação de frequência fundamental e energia.

Por outro lado, Eyben *et al.* (2010) apresenta o extrator *emo\_large* como uma composição de um conjunto extenso de extratores, incluindo MFCCs, coeficientes de codificação preditiva linear (LPC) e características de entropia espectral. Esses elementos são comumente aplicados para análise de emoções por sinal de voz. A combinação destes extratores engloba uma variedade de características fundamentadas na expressão emocional com base na voz.

Neste mesmo sentido, o extrator *emobase* é a composição dos extratores MFCCs, F0, coeficientes LPC, juntamente com características específicas baseadas na emissão dos sons da fala, como o acento e a entoação, bem como acentuação vocabular, também desenvolvido com foco na análise detalhada da emoção através da voz.

Como uma evolução desenvolvida para o INTERSPEECH 2010, o *emobase2010* é

uma nova versão do extrator emobase. O *emobase2010* se baseia em MFCCs, F0, formantes<sup>2</sup> e coeficientes LPC, adaptando-se às exigências específicas da competição. Já o conjunto *IS09 emotion* integra atributos como MFCCs, F0 e características temporais, projetado para tarefas específicas de identificação e análise de emoções na fala.

Baseado nas especificações do INTERSPEECH 2010 *Paralinguistic Challenge*, o *IS10 paraling* e *IS10 paraling compat* incorporam extratores como MFCCs, F0 e medidas de jitter e shimmer para análise paralinguística. A versão *compat* é uma implementação compacta que garante uma melhor consistência e interoperabilidade com diferentes sinais.

O conjunto *IS12 speaker trait* e *IS12 speaker trait compat* são extratores para a análise de características relacionadas a traços do locutor, como personalidade e estilo de fala, além dos formante. Formantes, neste caso, é um pico de energia em uma região do espectro sonoro, ele define vogais, timbre e outras propriedades sonoras. Este extrator enfatiza características relacionadas a traços individuais do locutor, como idade e gênero, e resposta em frequência do trato vocal. A versão *compat* é uma implementação compacta que garante a interoperabilidade com diferentes sinais.

O extrator *IS13 ComParE*, por sua vez, foi desenvolvido para a competição INTERSPEECH 2013 (SCHULLER *et al.*, 2013), abrangendo um amplo conjunto de características extraídas do sinal para análise paralinguística, e consolida diversos atributos do sinal sonoro, tais como: MFCCs, F0, Formantes, Entropia Espectral e Complexidade Espectral.

Finalmente, uma vez apresentadas todas as diferentes combinações de configurações de extratores produzidos pelo OpenSMILE, a Tabela 2 lista as combinações com seus respectivos número de atributos correspondentes.

---

<sup>2</sup> Os formantes atuam na definição do timbre de fontes sonoras. Embora o espectro de cada nota produzida por um instrumento possa apresentar variações significativas com a mudança de altura, as áreas associadas aos formantes permanecem consistentes, independentemente da frequência da nota. Assim, os formantes atuam como uma espécie de assinatura distintiva para uma fonte sonora específica.

Tabela 2 – Diferentes configurações de extratores produzidos pelo OpenSMILE e o número total de atributos.

Extrator de atributos	Número de atributos
avec2011	1942
avec2013	2268
emo_large	6553
emobase	989
emobase2010	1583
IS09_emotion	384
IS10_paraling	1582
IS10_paraling_compat (IS10_pc)	1582
IS11_speaker_state (IS11_ss)	4368
IS12_speaker_trait (IS12_st)	5757
IS12_speaker_trait_compat (IS012_stc)	6125
IS13_ComparE	6373

Fonte: Elaborada pelo autor.

Isto posto, vale destacar que a seleção deste conjunto de extratores específicos foram elencados devido ao seu destaque estabelecido na literatura e à sua eficácia comprovada em estudos anteriores para a classificação de doenças a partir de sinais de voz (ALMEIDA *et al.*, 2019; VAICIUKYNAS *et al.*, 2017). Portanto, quanto a estes extratores, visamos avaliar sua capacidade de extração de atributos em sinais de voz, afim de auxiliar na detecção da doença de Huntington através da análise de sinais de voz, que é descrita no Capítulo 5, deste modo estabelecer comparativos com relação a metodologia de extração e classificação proposta.

## 2.4 Modelos de aprendizado de máquina: abordagem clássica

Além da extração de atributos estão os diferentes modelos computacionais de classificação utilizados para obter informações com base nas amostras. Esta é uma fase em que busca-se desenvolver um modelo computacional capaz de separar de maneira satisfatória as amostras em suas respectivas classes, com base unicamente em suas construções matemáticas e nas informações contidas nos atributos extraídos das amostras, úteis ao treinamento e adaptação dos respectivos algoritmos de classificação.

Neste sentido, o K-Nearest Neighbors (KNN) é um algoritmo de classificação introduzido por Cover e Hart (1967). Atribui uma classe a um ponto não classificado com base na classe ou rótulo majoritário do conjunto de dados entre os seus  $k$  vizinhos mais próximos. Onde a proximidade é geralmente medida pela distância euclidiana. O KNN faz previsões atribuindo a classe mais frequente entre os  $k$  vizinhos mais próximos, medidos pela distância euclidiana.

Para uma instância de teste  $x$ , a predição  $y$  é dada por:

$$y = \arg \max_c \sum_{i=1}^k I(y_i = c) \quad (2.15)$$

onde  $y_i$  é a classe do vizinho mais próximo  $i$  de  $x$ .

Por sua vez o Random Forest (RF), descrito por Breiman (2001), é um classificador baseado na construção de um conjunto de árvores de decisão. Ele constrói várias árvores de forma independente e combina suas previsões para obter um resultado mais robusto. A previsão de um Random Forest é a média das previsões de suas árvores individuais:

$$f(x) = \frac{1}{n} \sum_{i=1}^n T_i(x) \quad (2.16)$$

onde  $T_i(x)$  é a decisão da  $i$ -ésima árvore.

Já o classificador ExtraTrees (ETrees) descrito por Geurts *et al.* (2006) é uma metodologia de aprendizagem automática baseada em árvores de decisão. De maneira semelhante ao classificador Random Forest, este classificador constrói várias árvores de decisão e toma decisões com base em uma média ponderada de suas saídas. O termo "extra" refere-se ao fato de que, durante a construção das árvores, amostras aleatórias são retiradas para cada divisão.

De outra forma, como apresenta Hastie *et al.* (2009), no contexto da aplicação de combinações lineares está a Análise Discriminante Linear (LDA) que representa uma técnica de redução de dimensionalidade para descobrir uma combinação linear das características originais do conjunto de dados. O objetivo é maximizar a razão das variâncias entre as classes em relação à variância dentro das classes. A LDA procura encontrar um vetor de projeção  $w$  que maximize a razão de dispersão entre classes e a dispersão dentro das classes:

$$J(w) = \frac{|w^T S_b w|}{w^T S_w w} \quad (2.17)$$

onde  $S_b$  é a matriz de dispersão entre classes e  $S_w$  é a matriz de dispersão dentro das classes.

Neste mesmo contexto estatístico, Hastie *et al.* (2009) descreve a Análise Discriminante Quadrática (QDA) como um método utilizado para classificação e discriminação entre múltiplas classes. Ao contrário da LDA, a QDA não assume a igualdade das matrizes de covariância entre as classes, permitindo modelar formas mais complexas de decisão. Suponha que tenhamos  $K$  classes e que  $n_k$  seja o número de observações na classe  $k$ . Denotamos por  $p$  o número de características (ou variáveis preditoras). Para uma classe específica  $k$ , a função de densidade de probabilidade multivariada é:

$$f_k(x) = \frac{1}{(2\pi)^{p/2} |\Sigma_k|^{1/2}} \exp\left(-\frac{1}{2}(x - \mu_k)^T \Sigma_k^{-1} (x - \mu_k)\right) \quad (2.18)$$

onde:  $x$  é um vetor de características de tamanho  $p$ ,  $\mu_k$  é o vetor médio das observações na classe  $k$ ,  $\Sigma_k$  é a matriz de covariância das observações na classe  $k$ , e  $|\Sigma_k|$  denota o determinante de  $\Sigma_k$ . A regra de classificação para a QDA é então determinada pela maximização da probabilidade a posteriori, dada pela fórmula de Bayes Hastie *et al.* (2009).

O classificador Gaussian Naive Bayes, descrito em Zhang (2004), é baseado no teorema de Bayes e assume que as características das amostras são condicionalmente independentes, dado o rótulo da classe. Ele é especialmente útil quando as características são contínuas e seguem uma distribuição normal. É um classificador que atribui pontos de dados à classe mais próxima, calculando distâncias com base na média e na variância dos dados. Incorpora métricas estatísticas relevantes para a análise de dados. Deste modo, para este classificador gaussiano, a probabilidade condicional é modelada como:

$$P(x_j|y) = \frac{1}{\sqrt{2\pi\sigma_{yj}^2}} \exp\left(-\frac{(x_j - \mu_{yj})^2}{2\sigma_{yj}^2}\right) \quad (2.19)$$

A decisão é feita usando o Teorema de Bayes.

Já o classificador Multinomial Naive Bayes, também inspirado no Teorema de Bayes e descrito por Schütze *et al.* (2008), é um modelo probabilístico que assume a independência condicional<sup>3</sup> entre as características dadas as classes. É especialmente útil para dados discretos, comumente aplicado em tarefas como classificação de texto. Dado um conjunto de características  $x = (x_1, x_2, \dots, x_d)$  e um conjunto de classes  $C = \{c_1, c_2, \dots, c_k\}$ , a probabilidade condicional de uma classe  $c_i$  para a amostra  $x$  é dada por:

$$P(c_i|x) = \frac{P(c_i) \cdot P(x_1|c_i) \cdot P(x_2|c_i) \cdot \dots \cdot P(x_d|c_i)}{P(x)} \quad (2.20)$$

A classificação, portanto, de uma nova instância  $x$  é realizada selecionando a classe que maximiza a probabilidade condicional, tal decisão também é baseada no Teorema de Bayes.

Por outro lado, baseado em uma abordagem de kernel está o *Support Vector Machine* (SVM), descrito por Vapnik (1999). O SVM é um poderoso algoritmo de aprendizado de máquina usado para classificação e regressão. A ideia principal por trás do SVM é encontrar

<sup>3</sup> A independência condicional é um conceito estatístico que descreve a relação entre eventos em um espaço amostral. Dois eventos são considerados condicionalmente independentes se a ocorrência ou não ocorrência de um evento não afeta a probabilidade de ocorrência do outro evento, dada uma terceira variável.

um hiperplano de separação ótimo que maximize a margem entre as classes. Dada uma coleção de dados de treinamento  $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ , onde  $x_i$  é um vetor de características e  $y_i$  é a classe associada, o objetivo do SVM é encontrar um vetor de pesos  $w$  e um termo de polarização  $b$  tal que o hiperplano de separação seja dado por  $w^T x + b = 0$ . Neste sentido, a função de decisão para uma nova instância  $x$  é dada por:

$$f(x) = \text{sign}(w^T x + b) \quad (2.21)$$

O SVM introduz o conceito de kernels para mapear os dados para um espaço de maior dimensionalidade, onde a separação linear pode ser alcançada. Um exemplo de kernel comum é o kernel de função de base radial (RBF):

$$K(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right) \quad (2.22)$$

onde  $\sigma$  é um parâmetro que controla a largura do kernel.

Por outro lado, baseado nos avanços de Friedman (2001), o Extreme Gradient Boosting (XGBoost) foi projetado como um algoritmo de aprendizado de máquina baseado em árvores de decisão que é conhecido por sua eficiência e rápida convergência na fase de treinamento. Ele pertence à categoria de métodos *ensemble*<sup>4</sup>, especificamente à classe de métodos de *boosting*<sup>5</sup>. O XGBoost otimiza a função objetivo regularizada que é uma combinação da função de custo e termos de regularização. A função objetivo geral do XGBoost é dada por:

$$\mathcal{L}(\theta) = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (2.23)$$

onde  $n$  é o número de amostras no conjunto de treinamento,  $l(y_i, \hat{y}_i)$  é a função de custo que mede a diferença entre a previsão ( $\hat{y}_i$ ) e o rótulo real ( $y_i$ ). Por fim, os  $\sum_{k=1}^K \Omega(f_k)$  são os termos de regularização para as  $K$  árvores no modelo.

Por fim, o classificador *Histogram-based Gradient Boosting* é uma técnica de aprendizado de máquina que pertence à família de algoritmos de *Gradient Boosting*. A principal diferença reside na forma como as árvores de decisão são construídas e atualizadas durante o treinamento, a implementação avaliada neste trabalho segue a descrição de Ke *et al.* (2017).

O objetivo é otimizar uma função objetivo regularizada, semelhante ao XGBoost. A função objetivo é definida como:

<sup>4</sup> Métodos ensemble são abordagens que combinam as previsões de vários modelos de aprendizado de máquina para melhorar o desempenho geral e a robustez do sistema.

<sup>5</sup> Os modelos são treinados de forma iterativa, onde cada novo modelo tenta corrigir os erros do modelo anterior. Pesos são atribuídos às instâncias de treinamento, dando mais importância às instâncias classificadas erroneamente.

$$\mathcal{L}(\theta) = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (2.24)$$

onde  $l$  é a função de custo,  $\hat{y}_i$  é a previsão,  $f_k$  são as árvores e  $\Omega$  é a penalidade de regularização.

## 2.5 Redes Neurais Artificiais

Redes Neurais Artificiais (RNA) são modelos computacionais inspirados no funcionamento do sistema nervoso biológico, projetados para realizar tarefas de aprendizado e reconhecimento de padrões. Uma unidade fundamental em redes neurais é o neurônio, que pode ser modelado matematicamente como uma combinação linear entre pesos e valores que correspondem ao sinal de entrada. Duas formulações iniciais incluem o Perceptron e o neurônio de McCulloch-Pitts (ROSENBLATT, 1958; MCCULLOCH; PITTS, 1943), que serviram como base para o desenvolvimento de redes neurais mais complexas.

O treinamento de uma rede neural envolve ajustar os pesos, representados pelo vetor  $w$ , de modo que a saída da rede se aproxime dos valores desejados. Esse procedimento toma como base algoritmos de retropropagação do erro, onde o ajuste dos pesos são feitos visando a redução do erro obtido na saída, de forma iterativa. Isso é geralmente realizado utilizando algoritmos de otimização, como o gradiente descendente.

### 2.5.1 Multilayer Perceptron (MLP)

O Perceptron de Múltiplas Camadas (MLP, de *MultiLayer Perceptron* em inglês) é uma das estruturas de RNA mais empregadas, sendo considerado uma evolução natural do Perceptron. A principal limitação do Perceptron é sua capacidade de aprender a discriminar apenas funções que são separáveis linearmente, enquanto os MLPs têm a habilidade de aprender funções não-linearmente separáveis, o que os torna bastante potentes e flexíveis (SIMON; HAYKIN, 2009). Uma representação de sua saída pode ser vista na Equação 2.25.

$$\hat{y} = \phi \left( \sum_{i=1}^p x_i w_i + b \right) \quad (2.25)$$

onde  $\phi$  corresponde à função de ativação,  $x_i$  é o vetor de entrada e  $w_i$  é o vetor de pesos que modula a entrada, ambos possuindo a mesma dimensão  $p$ . O termo de viés  $b$ , bem como o vetor  $w_i$ , são parâmetros que são ajustados durante o processo de treinamento do Perceptron. Este

treinamento ocorre de forma iterativa, como indicado pela Equação 2.26,

$$w_i(t+1) = w_i(t) + \eta(y - \hat{y}(t))x_i, \quad \forall i = 1, \dots, p \quad (2.26)$$

onde  $t$  representa a iteração de treinamento,  $y$  é a classe da amostra e  $\hat{y}(t)$  é a previsão realizada para esta amostra na respectiva iteração.

Já na fase de treinamento, o termo  $\eta$  na Equação 2.26 refere-se a taxa de aprendizado, que tem como função ponderar o erro de previsão. As limitações sobre a capacidade para problemas não-lineares motivaram o uso de funções que pudessem contornar este comportamento da saída da rede. Neste sentido uma das funções aderentes é a sigmoide, encarregada de introduzir a não linearidade no vetor de características. Cada neurônio, portanto, tem a sua saída determinada pela camada na qual se encontra, conforme apresentado na Equação 2.27,

$$\hat{y}_{l_k} = \begin{cases} \phi_{0_k} \left( \sum_{i=1}^p x_i w_{0_{k_i}} + b_{0_k} \right) & \text{se } l = 0 \\ \phi_{l_k} \left( \sum_{i=1}^D \hat{y}_{l-1} w_{l_{k_i}} + b_{l_k} \right) & \text{se } l \neq 0 \end{cases} \quad (2.27)$$

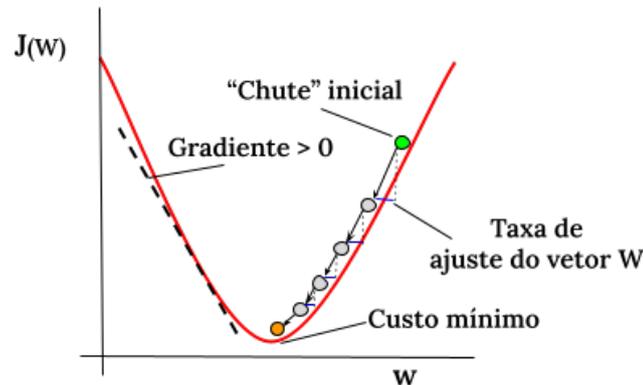
onde  $l$  representa o número da camada e  $k$  é a posição do neurônio nessa camada, sendo  $x$  um vetor de atributos de dimensão  $p$  e  $\phi$  é a função de ativação não linear.

O treinamento de um MLP normalmente consiste na propagação direta dos dados de entrada pela rede para produzir uma saída, seguida pela retropropagação do erro pela rede. Esse processo é conhecido como *backpropagation* e é comumente realizado em conjunto com um algoritmo de otimização, como o Descida de Gradiente (RUMELHART *et al.*, 1986).

Considere a função  $J$  como a representação da taxa de erros, a Figura 3 ilustra o sentido da adaptação do vetor  $w$  visando a minimização de  $J$ , esta adaptação é direcionada pelo vetor gradiente, tal adaptação é proporcional aos ajustes inputado pela taxa de aprendizado  $\eta$ .

Apesar das vantagens dos MLPs, eles têm suas limitações. Uma delas é o problema de *vanishing gradients*, onde os gradientes se tornam muito pequenos à medida que são propagados de volta através das camadas da rede durante o treinamento. Isso torna o treinamento de MLPs profundos (i.e., com muitas camadas ocultas) desafiador. Novas técnicas e arquiteturas foram propostas para resolver este problema, como a inicialização de pesos cuidadosa e as unidades de normalização em lote (HE *et al.*, 2015; IOFFE; SZEGEDY, 2015).

Figura 3 – Função de custo  $J$  de  $w$  apresentando a descida do gradiente e taxa de ajuste do vetor  $W$ . Em verde, a configuração de pesos iniciais, em laranja, a configuração desejada.



Fonte: Elaborada pelo autor.

### 2.5.1.1 O problema do Vanishing Gradient

Conforme apresentado na seção anterior, o treinamento de RNAs mais profundas onde ocorre uma maior retropropagação do erro pode incorrer em um comportamento não desejável conhecido como "*vanishing gradient*". Segundo He *et al.* (2016), este fenômeno caracteriza-se como uma das dificuldades encontradas durante o treinamento de redes neurais profundas. Ele ocorre quando os gradientes das funções de custo em relação aos parâmetros (pesos) das camadas mais profundas da rede se aproximam de zero durante a retropropagação, resultando em uma atualização insignificante dos pesos e, conseqüentemente, em uma aprendizagem ineficaz.

Para mitigar o problema do *vanishing gradient*, são adotadas várias estratégias, tais como a escolha de funções de ativação mais apropriadas no sentido de que a magnitude de sua derivada não seja próxima a zero, o uso de inicializações de peso adequadas, evitando valores de gradientes reduzidos já no início do treino, o que pode resultar em uma convergência precoce, outra alternativa é a normalização de lotes (operação conhecida também como *batch normalization*), além disso, mais recentemente, arquiteturas de rede mais avançadas, como redes residuais (ResNets) apresentam resultados promissores para este problema (HE *et al.*, 2016).

### 2.5.2 Rede neural residual (ResNet)

As Redes Neurais Residuais (ResNets), descritas em He *et al.* (2016), foram propostas para abordar desafios encontrados em treinamentos profundos, tal como o problema do *vanishing gradient*. A ideia fundamental por trás das ResNets é a introdução de atalhos

ou conexões residuais que pulam uma ou mais camadas, permitindo um fluxo mais direto do gradiente durante a retropropagação, o termo resíduo, neste caso, refere-se as diferenças entre a entrada e a saída desejada.

Seja  $x$  a entrada para uma camada, e seja  $F(x)$  a saída desejada dessa camada após a aplicação de transformações não lineares. A operação residual é definida como:

$$\text{Residual}(x) = F(x) + x \quad (2.28)$$

A saída final da camada é então  $y = \text{Residual}(x)$ . Essa formulação permite que o modelo aprenda a ajustar os pesos para reduzir a saída residual para zero.

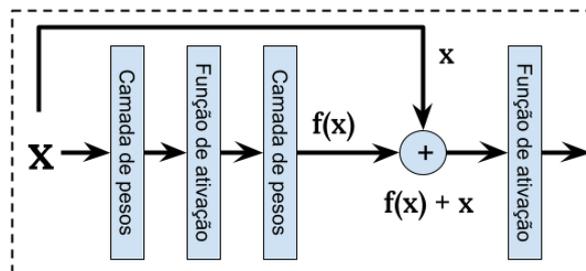
O bloco residual proposto neste tipo de arquitetura é a unidade básica em uma ResNet, consistindo em duas camadas convolucionais (ou outras operações) e uma operação de adição do atalho. Seja  $H(x)$  a saída da primeira camada convolucional e  $F(x)$  a saída da segunda camada convolucional, então a saída do bloco residual é dada por:

$$y = \text{Residual}(x) = F(H(x)) + x \quad (2.29)$$

Em um bloco regular de uma rede neural profunda, após a combinação linear que acontece na camada de pesos, o potencial de ativação produzido é direcionado a função de ativação para produzir a saída da camada. Por outro lado, a Figura 4 apresenta a principal diferença imposta pelo bloco residual. Após a saída da camada de pesos, o resultado denotado por  $f(x)$  é combinado com os valores da entrada da camada, via conexão residual originária do vetor de dados na entrada.

Tal combinação produz o mapeamento residual  $f(x) + x$ , que é como o bloco residual deriva seu nome, só então o potencial de ativação é direcionado a função de ativação. Esta mudança modifica o potencial de ativação na tentativa de que a saída da função não seja um valor próximo a zero, afim de que não haja a atenuação e anulação do gradiente na retropropagação.

Figura 4 – Bloco Residual.



Fonte: Elaborada pelo autor.

### 2.5.3 Redes Neurais Convolucionais

A convolução apresenta-se a aplicação de um filtro com o propósito de identificar e extrair o maior número de informações relevantes em uma determinada partição da imagem ou do sinal de entrada, ou como detalha Albawi *et al.* (2017), é uma operação fundamental para extrair dados das camadas anteriores como é o caso das Redes Neurais Convolucionais (CNN). Conforme a arquitetura da CNN, diferentes filtros são produzidos e aplicados com o objetivo de extrair características relevantes tais como detectores passa alta, identificando assim transições de borda, gradientes bem como contrastes e texturas, no caso de imagens. Mas também são úteis como filtro passa baixa, na detecção de borrimento, regiões de atenuação do sinal e, no caso de imagens, áreas de texturas uniformes.

#### 2.5.3.1 Operação de Convolução

Para funções discretas, a convolução de duas funções  $f(t)$  e  $g(t)$ , onde  $t$  representa o tempo é expressa como uma soma discreta:

$$h[n] = (f * g)[n] = \sum_{k=-\infty}^{\infty} f[k]g[n-k] \quad (2.30)$$

onde  $n$  é uma variável discreta que representa o índice do sinal discreto. Essa equação reflete a soma dos produtos dos elementos correspondentes de  $f$  e  $g$ , deslocados uns em relação aos outros. A propriedade comutativa da convolução afirma que  $f * g = g * f$ , o que significa que a ordem das funções na operação não afeta o resultado (BURRUS; PARKS, 1985; RHODES, 1981).

Na convolução em amostras espaciais de duas dimensões, como é o caso de imagens, o filtro  $K$  em um sinal bidimensional  $S$ , o filtro percorre o sinal por completo. Quando existe sobreposição entre ambos, é realizado um somatório dos produtos de elementos correspondentes, conforme a Equação 2.31:

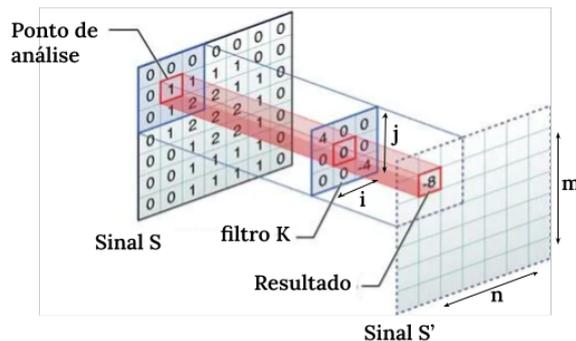
$$S'(i, j) = (S * K)(i, j) = \sum_m \sum_n S(m, n)K(i - m, j - n) \quad (2.31)$$

onde  $S'$  é o mapa de características gerado pela convolução do filtro  $K$  no sinal  $S$ . Na figura 5 é ilustrada a operação de convolução, no exemplo, um filtro  $K$  de dimensões  $ixj$  é aplicado a uma matriz  $S$  de dimensões  $mxn$ .

No exemplo da Figura 5, a região de análise do sinal  $S$  apresenta nove valores, com um ponto central de análise, no contexto de imagem, este ponto é o *pixel* central da convolução.

O filtro  $K$  é aplicado com base na sobreposição ponto a ponto com os valores do sinal  $S$ , as operações de multiplicação e soma é executada, produzindo o novo valor  $-8$ , que por sua vez é armazenado na mesma posição do ponto central de análise, porém, na matriz de saída da operação, sinal  $S'$ .

Figura 5 – Operação de convolução em um sinal bidimensional  $S$  sob um filtro convolucional  $K$ .



Fonte: Elaborada pelo autor.

Além de modelos tradicionais para a classificação das amostras, este estudo também considera modelos mais recentes baseados em redes neurais profundas. Neste sentido, esta seção apresenta uma série de arquiteturas que utilizam como entrada de dados uma estrutura tensorial que representa a amostra e a partir de então, executam suas camadas internas a seleção de melhores parâmetros para cada conjunto de neurônios afim de melhor classificar amostras novas amostras.

Além das ResNet, descritas na seção de redes residuais, existem arquiteturas recentes, com foco na redução do custo computacional, propondo arquiteturas sintéticas, úteis para execução em dispositivos embarcados e dispositivos móveis, onde os recursos computacionais são geralmente escassos. Neste sentido, elencamos nas seções a seguir, arquiteturas que se destacam, dentre as outras, pelo seu projeto com foco neste cenário de aplicações.

### 2.5.3.2 Arquitetura MnasNet:

Dentro deste contexto, visando a construção de uma RNA convolucional com redução de custo computacional, Tan *et al.* (2018) introduziram a rede neural MnasNet para aplicações em tempo real com foco em plataformas com recursos limitados. A MnasNet, abreviação de *Mobile Neural Architecture Search*, é uma arquitetura otimizada para dispositivos móveis, projetada através de um processo de busca automática de arquiteturas neurais para atender às restrições de

recursos presentes em dispositivos móveis.

MnasNet utiliza a técnica de busca de arquitetura neural com restrições de recursos, considerando fatores como tamanho e poder computacional limitados em dispositivos móveis. A arquitetura resultante é eficiente e leve, usando camadas como operações factorizadas e convoluções separáveis para reduzir a carga computacional.

#### 2.5.3.3 *Arquitetura MobileNet V2:*

MobileNet V2, concebido por Sandler *et al.* (2018), foi adaptada para funcionar em dispositivos móveis com recursos de hardware limitados. MobileNetV2 é uma evolução do MobileNet original, projetado para fornecer eficiência e desempenho para tarefas de visão computacional em dispositivos móveis. Incorpora blocos residuais, normalização por lote (*batch normalization*) e ativação linear por peça (ReLU) para melhorar a capacidade representacional. Além disso, utiliza uma técnica chamada "inverted residuals" para otimizar o fluxo de informações dentro da rede.

#### 2.5.3.4 *Arquitetura ResNeXt:*

Já a arquitetura ResNeXt, conforme apresentado por Xie *et al.* (2017), é uma arquitetura de rede neural modularizada concebida para a classificação de padrões. Sua proposta é uma construção com menos camadas para reduzir os custos computacionais e introduz o conceito de cardinalidade como uma dimensão adicional para melhorar a precisão. A ResNeXt é uma extensão da arquitetura ResNet que introduz o conceito de cardinalidade para melhorar a representação.

#### 2.5.3.5 *Arquitetura ShuffleNet V2:*

A ShuffleNet V2 é uma arquitetura de rede neural proposta por Ma *et al.* (2018) para resolver tarefas de classificação de padrões. Ela representa uma evolução direta da ShuffleNet (ZHANG *et al.*, 2018) original e foi projetada para melhorar a eficiência em termos de velocidade e precisão. ShuffleNet V2 é uma arquitetura projetada para otimizar a eficiência computacional, principalmente através da utilização de operações de convolução que reduzem o custo computacional.

Apresenta a técnica de "shuffle" para realizar convoluções de ponto flutuante eficientemente.

entes, minimizando a sobrecarga computacional. Utiliza blocos de construção como unidades Shuffle e camadas de convolução especializadas para otimizar ainda mais o desempenho em dispositivos com recursos limitados.

#### 2.5.3.6 *Arquitetura Wide ResNet:*

A *Residual Wide Network*, proposta por Zagoruyko e Komodakis (2016), é uma arquitetura de rede neural profunda conhecida pela sua eficiência e desempenho. A Wide ResNet é uma extensão da arquitetura ResNet, onde a largura dos blocos residuais é aumentada para melhorar a capacidade de aprendizado.

Ao aumentar o número de canais em cada camada, a Wide ResNet visa melhorar a representação e a discriminação de características. Isso permite que a rede aprenda representações mais ricas, tornando-a especialmente eficaz para tarefas de classificação em conjuntos de dados desafiadores.

## 2.6 Avaliação de desempenho

Nesta seção apresentamos as métricas que balizarão a metodologia quanto a avaliação de desempenho do algoritmo de descrição de sinais, proposto neste trabalho. Esta seção se divide em duas partes, a primeira abrange a métrica estatística para discriminação das amostras entre e intra classe. A segunda parte, abrange as métricas clássicas quanto ao modelos de classificação.

### 2.6.1 *Análise de Variância (ANOVA) quanto a discriminação de classes*

A Análise de Variância (ANOVA), originalmente descrita por Sthle *et al.* (1989), é uma técnica estatística utilizada para comparar médias entre grupos independentes. Essa metodologia é valiosa quando se deseja determinar se há diferenças significativas nas médias de variáveis dependentes entre diferentes níveis de uma variável categórica. ANOVA avalia a variação total nos dados e a divide em componentes atribuídas à variação entre grupos e à variação dentro dos grupos.

A formulação para um modelo de ANOVA com um fator pode ser expressa como:

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij} \quad (2.32)$$

onde  $Y_{ij}$  é a observação na  $j$ -ésima unidade experimental do  $i$ -ésimo grupo, já o  $\mu$  é a média global dos dados,  $\alpha_i$  é o efeito do  $i$ -ésimo grupo ( $i = 1, 2, \dots, k$ ) e  $\varepsilon_{ij}$  é o erro aleatório associado

à  $j$ -ésima observação no  $i$ -ésimo grupo.

A hipótese nula ( $H_0$ ) da ANOVA afirma que não há diferenças significativas entre as médias dos grupos, enquanto a hipótese alternativa ( $H_1$ ) sugere que pelo menos uma média é significativamente diferente. Deste modo, podemos definir a estatística de teste da ANOVA como sendo baseada na razão das variações entre grupos e dentro dos grupos.

#### 2.6.1.1 ANOVA e o Valor-p para algoritmo descritores de características

Sabe-se que o valor-p desempenha um papel crucial na tomada de decisões estatísticas. O valor-p é a probabilidade de observar uma estatística de teste tão extrema quanto a observada, assumindo que a hipótese nula é verdadeira. Se o valor-p for inferior a um determinado nível de significância (por exemplo, 0,05), rejeita-se a hipótese nula em favor da hipótese alternativa. Pode-se dizer que  $H_0$  denota que não há diferenças significativas entre as médias dos grupos, já o  $H_1$  indica que pelo menos uma média é significativamente diferente. Se o valor-p é pequeno, isso sugere que a variação entre os grupos é significativa em comparação com a variação dentro dos grupos, levando à rejeição da hipótese nula.

Vale destacar que, em geral, ANOVA é aplicado a situações quando se tem mais de dois grupos populacionais, contudo, conforme discute Mishra *et al.* (2019), o teste ANOVA não se limita apenas a esta condição. Pode-se dizer que quando ANOVA é aplicado a dois grupos de amostras, os resultados mostram-se similares ao teste T de Student (*T* Teste) para duas populações, sendo este considerado um tipo especial de ANOVA para duas populações.

Portanto, se a ANOVA indica diferenças significativas entre as médias, isso sugere que o algoritmo de extração de atributos tem a capacidade de discriminar efetivamente entre os grupos. Essa informação é valiosa em problemas onde a distinção entre grupos é fundamental. A análise ANOVA já foi explorada para avaliar a discriminação e seleção de atributos em trabalhos anteriores (JOHNSON; SYNOVEC, 2002; NASIRI *et al.*, 2022; PINEDA *et al.*, 2020).

#### 2.6.2 Métricas quanto ao Desempenho em Modelos de Classificação

A avaliação de desempenho em modelos de classificação é uma etapa crucial na análise e seleção de algoritmos. Neste sentido, para avaliar a metodologia do algoritmo descritor apresentado neste trabalho bem como seu desempenho quando combinado com algoritmos de classificação, cabe avaliar diversas métricas visando mensurar diferentes aspectos quanto ao desempenho dos modelos.

A matriz de confusão é uma tabela que resume o desempenho de um modelo de classificação, discriminando as instâncias classificadas corretamente e aquelas classificadas incorretamente em termos de suas verdadeiras classes, representada pela matriz da Figura 6:

Figura 6 – Matriz de confusão, apresentando os diferentes valores de verdadeiros positivos, falsos positivos, falsos negativos e verdadeiros negativos.

		Valor Predito	
		Sim	Não
Real	Sim	Verdadeiro Positivo (TP)	Falso Negativo (FN)
	Não	Falso Positivo (FP)	Verdadeiro Negativo (TN)

Fonte: Elaborada pelo autor.

No exemplo da figura, na diagonal principal, temos os casos onde o valor de classificação esperado (Valor Real) coincide com o valor de classificação predito pelo modelo (Valor Predito). Nas demais células, temos os casos onde houve uma indicação errada quanto a classe da amostra, representando os falsos negativos e falsos positivos.

Para conduzir a comparação, este estudo empregará métricas de avaliação derivadas de uma matriz de confusão. Esta matriz será construída comparando as amostras validadas por um especialista com os resultados produzidos por cada método de classificação empregado. A matriz de confusão produzirá valores como Verdadeiros Positivos (TP), Verdadeiros Negativos (TN), Falsos Positivos (FP) e Falsos Negativos (FN), que servirão posteriormente como base para o cálculo de várias métricas de avaliação, como descreve Vásquez-Correa *et al.* (2021).

Em problemas de classificação, onde pretendemos rotular instâncias às classes, existem diversas métricas de desempenho usadas para avaliar o desempenho do modelo de classificação. Neste trabalho, sugerimos as métricas clássicas para compor a avaliação: Acurácia, Precisão, Sensibilidade, Recall, Especificidade, F1-Score e AUC.

A acurácia é a métrica mais direta e é definida como a proporção de instâncias previstas corretamente em relação ao total de instâncias no conjunto de dados, de acordo com a Equação 2.33:

$$Acc = \frac{TP + TN}{TP + FN + FP + TN} \quad (2.33)$$

A precisão concentra-se na exatidão das previsões positivas feitas pelo modelo. De acordo com a Equação 2.34:

$$Prec = \frac{TP}{TP + FP} \quad (2.34)$$

Recall é o mesmo que sensibilidade. Representa a capacidade do modelo de identificar todas as instâncias relevantes. De acordo com a Equação 2.35:

$$Recall = \frac{TP}{TP + FN} \quad (2.35)$$

O F1-Score é a média harmônica de precisão e recuperação. Ele equilibra precisão e recall. De acordo com a Equação 2.36:

$$F1 - Score = \frac{2 \cdot Prec \cdot Recall}{Prec + Recall} \quad (2.36)$$

A Especificidade, ou Taxa de Verdadeiros Negativos, mede a capacidade do modelo de identificar corretamente as instâncias negativas. É calculada como o quociente entre os verdadeiros negativos e a soma dos verdadeiros negativos e falsos positivos:

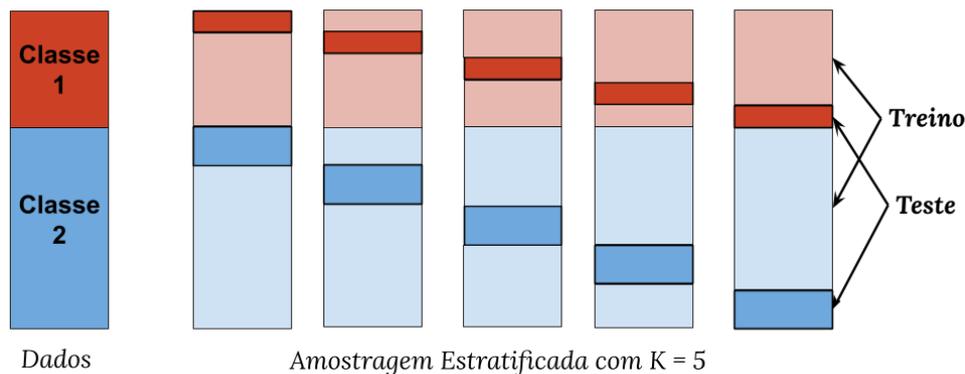
$$Esp = \frac{TN}{TN + FP} \quad (2.37)$$

A especificidade é relevante quando a ênfase está na identificação correta das instâncias negativas.

Por outro lado, a Área sob a curva ROC (AUC) é uma métrica de desempenho que avalia a capacidade discriminativa de um modelo em diferentes pontos de corte. Representa a área sob a curva ROC, que mostra a taxa de TP em função da taxa de FP. Uma AUC maior indica um melhor desempenho discriminativo do modelo.

Neste trabalho, todos as aplicações consideraram na fase de classificação dos tensores espectrais a estratégia de validação cruzada com amostragem estratificada, Figura 7. Esta metodologia procura garantir que cada dobra seja representativa de todos os estratos dos dados, de modo que cada classe seja aproximadamente representada igualmente em cada dobra de teste. A metodologia tem sido utilizada em trabalhos recentes da literatura para a tarefa de classificação de doenças (GANSER *et al.*, 2021; GAMBOA *et al.*, 2022).

Figura 7 – Exemplo visual da metodologia de validação cruzada com amostragem estratificada.



Fonte: Elaborada pelo autor.

### **2.6.3 Conclusão do Capítulo**

No encerramento deste abrangente capítulo de fundamentação teórica, exploramos conceitos fundamentais que fornecem alicerces sólidos para a pesquisa proposta. A Teoria dos Grafos apresentou uma base matemática para representar e analisar relações complexas, destacando a importância de diferentes matrizes na descrição de grafos direcionados. A DFT foi apresentada, evidenciando sua utilidade na conversão de sinais para o domínio da frequência, essencial para a robustez no processamento de sinais, especialmente diante de ruídos. A discussão sobre a extração de atributos de sinais introduziu métodos clássicos e especializados em voz.

Encerramos este capítulo explorando modelos de classificação tradicionais além de redes neurais profundas, incluindo, na última seção as métricas de desempenho quanto a utilização da metodologia para classificação de amostras. Delineando um arcabouço abrangente das técnicas que serão empregadas na construção da metodologia e avaliação dos resultados.

### 3 DESCRITOR DE CARACTERÍSTICAS EM COMPONENTES ESPECTRAIS BASEADO EM DIGRAFO: A METODOLOGIA

A metodologia proposta neste capítulo define o processamento de um sinal, ou conjuntos de sinais, com ênfase na construção de uma representação única consolidada na forma de um grafo direcionado, com arestas ponderadas. Neste intuito, a tarefa vai desde a conversão do sinal para o domínio da frequência, onde a análise é focada nas componentes espectrais até a construção de tensores espectrais, cuja representação é proposta sob três diferentes perspectivas.

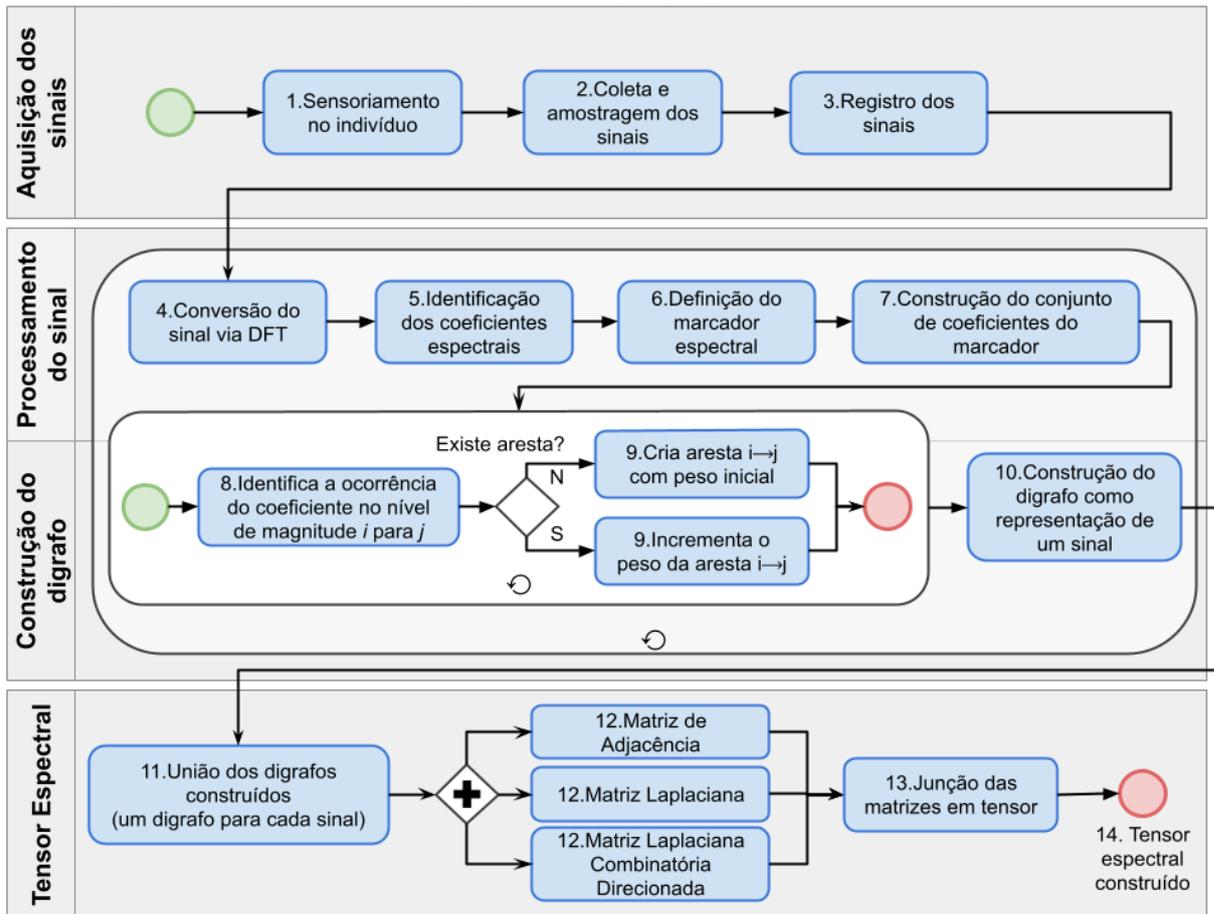
A Figura 8 apresenta o fluxograma que ilustra a proposta do descritor de sinal apresentado neste capítulo. Este fluxo sumariza de maneira mais abrangente as etapas que iniciam desde a produção do sinal biológico pelo paciente, coleta e registro do sinal (Etapas 1-3); passando pela etapa de conversão para o domínio da frequência e identificação dos coeficientes espectrais com base no marcador proposto (Etapas 4-7); seguido da construção das arestas ponderadas (Etapas 8-9); Por fim, união dos digrafos (cenário multicanal), produção das matrizes e consolidação no tensor espectral (Etapas 10-13). Nas seções a seguir são apresentadas cada uma destas etapas propostas em maiores detalhes.

#### 3.1 Domínio da frequência

A representação do sinal no domínio da frequência confere uma melhor abordagem para encontrar padrões de frequências e suas respectivas magnitudes, o que possibilita uma análise do grau de importância de cada componente no sinal original. Em se tratando de sinais biológicos, tal como EEG, podemos dizer que tal propriedade mostra-se relevante para a distinção de diferentes faixas de frequência, como *delta* (1~3Hz), *theta* (4~7Hz), *alpha* (8~12Hz), *beta* (12~28Hz) e *gamma* (30~200Hz). Frequências que podem ser associadas a diferentes estados mentais, variando desde uma mente sonolenta, passando por relaxamento ou até situações de alerta, além de representar processos cognitivos (SIJLY *et al.*, 2016; TATUM, 2021).

Entretanto, a abordagem no domínio da frequência apresenta alguns desafios: o sinal originalmente possui ruído embutido que aparece em parte ou em toda a amostra, sendo necessário mecanismos para atenuar a influência do ruído. Quanto as frequências dominantes no domínio do tempo, os componentes espectrais e a distribuição de energia do sinal não são facilmente identificáveis, contudo operações em componentes espectrais são mais eficientes, já que o sinal está decomposto em suas frequências formadoras.

Figura 8 – Fluxo da metodologia de extração proposta.



Fonte: Elaborada pelo autor.

### 3.2 Encontrando pontos marcadores nas componentes espectrais do sinal

Neste sentido, considerando a abordagem espectral, é proposto o processamento das amplitudes das componentes de frequência. Esta abordagem sugere a identificação de pontos relevantes em relação as componentes de frequência. Tais pontos de análise são chamados nesta proposta de **marcadores de componentes espectrais**. Os marcadores de componentes espectrais possuem magnitudes e, com base em seu valor e modo como surgem no sinal, são identificados ao longo da amostra na sua forma espectral.

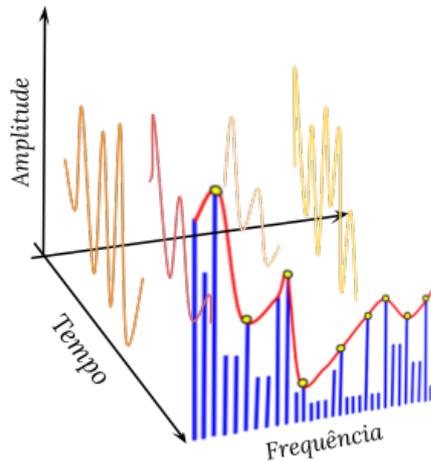
Deste modo, seja  $f$  a função que representa o sinal de entrada, a DFT atua como meio para produzir o domínio da função  $\phi$  e representa os componentes de frequência e sua magnitude:  $f(x) \xrightarrow{\mathcal{F}} \phi(k)$ . Assim, o cálculo sob a perspectiva discreta é apresentado como uma combinação utilizando a soma do produto de cada ponto no sinal original. Logo, aplicando DFT

ao sinal  $f$  temos:

$$\phi(k) = \sum_{n=0}^{N-1} f(x) \cdot e^{-j\frac{2\pi}{N}kn} \quad (3.1)$$

onde  $N$  é o número total de amostras,  $\phi(k)$  significa k-ésimo componente da DFT,  $j$  é a unidade imaginária e  $f(x)$  são os valores discretos do sinal no domínio do tempo. A Figura 9 ilustra a relação entre os domínios e apresenta o sinal sob as duas perspectivas, no tempo (sinal em laranja) e na frequência (sinal em azul), com as frequências decompostas (outras cores). No domínio da frequência, os componentes espectrais são representados por barras azuis, já os pontos marcadores nas componentes espectrais são apresentados nos pontos amarelos.

Figura 9 – Representação do sinal sob o domínio do tempo e frequência.



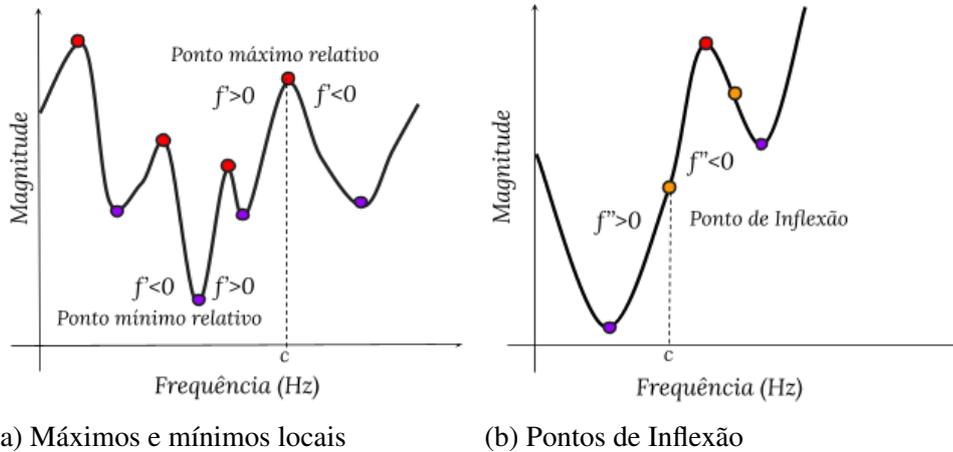
Fonte: Elaborada pelo autor.

Os componentes espectrais de maior magnitude na DFT, por exemplo, representam as frequências predominantes no sinal analisado, isto é, se estivermos analisando um sinal de áudio de uma música, os componentes espectrais de maior magnitude podem corresponder às notas musicais mais altas da música. Assim, o estudo de componentes espectrais têm sido explorados em diversos cenários, entre eles para classificar doenças (ALTAHERI *et al.*, 2023). Este processo de conversão do sinal via DFT, detalha, portanto, a etapa 4 da Figura 8.

### 3.2.1 Análise proposta dos extremos locais relativos nas componentes espectrais da função $\phi$

Seja  $\phi$  uma função definida em um intervalo  $I$  e sejam  $x_1$  e  $x_2$  quaisquer dois pontos em  $I$ . Se  $\phi(x_1) < \phi(x_2)$  quando  $x_1 < x_2$  então  $\phi$  está aumentando em  $I$ . Por outro lado, se  $\phi(x_2) < \phi(x_1)$  quando  $x_1 < x_2$ , então  $\phi$  está diminuindo em  $I$ .

Figura 10 – Análise proposta dos extremos locais relativos da função  $\phi$ .



Fonte: Elaborada pelo autor.

Conforme apresenta Leithold (1994), suponha que  $\phi$  seja contínua em  $I = [a, b]$  e diferenciável em  $I' = ]a, b[$ , de acordo com *teste da primeira derivada*: Se  $\phi'(x) > 0$  em qualquer ponto  $x \in I'$ , então  $\phi$  está aumentando em  $I$ . Se  $\phi'(x) < 0$  em qualquer ponto  $x \in I'$ , então  $\phi$  está diminuindo em  $I$ . De acordo com a mudança de sinal da função  $\phi$ , o ponto onde ocorre a mudança de sinal é denominado extremo local relativo.

Suponha que  $c$  seja um extremo local relativo da função contínua  $\phi$  e que  $c$  pertença ao intervalo  $I$ , se  $\phi'$  muda de negativo para positivo em  $c$ , então  $\phi$  tem um ponto mínimo relativo em  $c$ . Se  $\phi'$  mudar de positivo para negativo em  $c$ , então  $\phi$  terá um ponto máximo relativo em  $c$ . Se  $\phi'$  não muda de sinal em  $c$ , então  $c$  não é um extremo local de  $\phi$ .

A Figura 10a ilustra o gráfico da função  $\phi$ , no plano bidimensional das frequências e magnitudes em  $\phi$ . Quando o sinal da primeira derivada muda, há um ponto de máximo ou mínimo relativo local. Neste exemplo, a mudança de polaridade acontece no ponto  $c$ , sendo este um ponto relativo local, se é um mínimo ou máximo, dependerá dos pontos adjacentes.

### 3.2.2 Ponto de inflexão nas componentes espectrais da função $\phi$

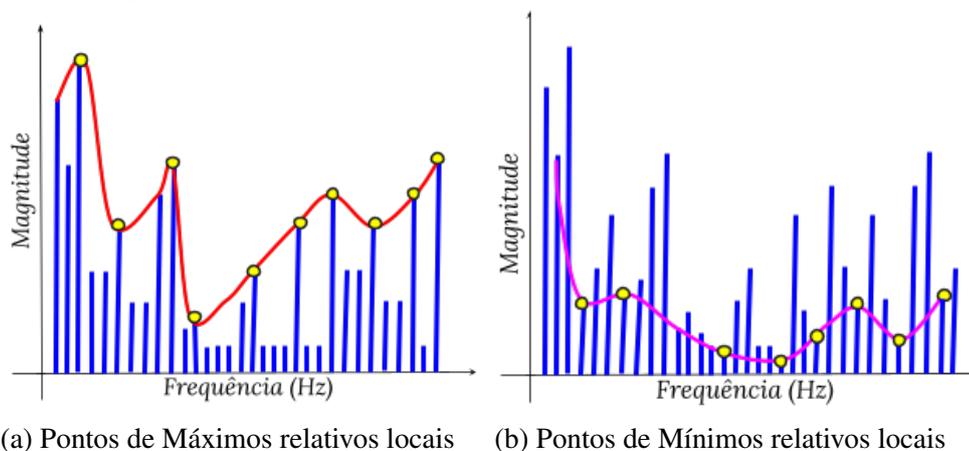
Seja  $\phi$  uma função definida em um intervalo  $I$ , suponha que  $\phi$  seja contínua em  $I = [a, b]$  e diferenciável em  $I' = ]a, b[$ , como descreve Leithold (1994), o *teste da segunda derivada* é dado por: Se  $\phi''(x) > 0$  em qualquer ponto  $x \in I'$ , quando  $x < c$ , e  $\phi''(x) < 0$  em qualquer ponto  $x \in I'$ , quando  $x > c$ , de acordo com a mudança de sinal da função  $\phi$ , o ponto onde ocorre a mudança de sinal é chamado de ponto de inflexão da função. A Figura 10b ilustra o gráfico da função  $\phi$  em um plano bidimensional das frequências por suas magnitudes, quando

o sinal da segunda derivada muda, há um ponto de inflexão no ponto  $c$ .

### 3.2.2.1 O marcador espectral alpha ( $\alpha$ ) proposto

A Figura 11a apresenta o sinal  $\phi$  destacando componentes espectrais específicos. Na abordagem que é proposta, esta metodologia define pontos que representam máximos relativos locais, aqui também chamamos de coeficientes de magnitude máxima local, cujo conjunto define o marcador espectral alpha ( $\alpha$ ) proposto. Este marcador espectral indica onde a função  $\phi$  atinge valores máximos dentro de uma vizinhança, são pontos críticos e identificam picos locais ou altos locais em uma curva. Tal marcador indica as frequências que possuem valores máximos locais, em detrimento de frequências com menor magnitude. Tal informação é útil para identificar frequências que contribuem com maior energia para a composição do sinal original, conforme a magnitude das componentes espectrais. Deste modo, estabelecemos a proposta do marcador espectral  $\alpha$  como um conjunto de pontos de máximos relativos locais.

Figura 11 – Marcadores espectrais alfa ( $\alpha$ ) e beta ( $\beta$ ) propostos no sinal da função  $\phi$ .



Fonte: Elaborada pelo autor.

Na análise de sinais de voz, por exemplo, os pontos de máximo relativos podem corresponder a características específicas do som, como picos de intensidade vocal ou pontos de transição entre diferentes fonemas. Por outro lado, em sinais multicanais tais como sinais EEG, esses pontos podem representar variações abruptas na atividade cerebral, como picos de elétricos associados a eventos cognitivos ou emocionais. (SIJLY *et al.*, 2016; TATUM, 2021)

### 3.2.2.2 O marcador espectral beta ( $\beta$ ) proposto

Por outro lado, os componentes espectrais de menor magnitude na DFT representam frequências com uma contribuição menos significativa para a composição do sinal. A Figura 11b apresenta o sinal  $\phi$  destacando componentes espectrais que representam mínimos relativos locais, o conjunto formado por estes pontos define o marcador espectral beta ( $\beta$ ) proposto. A proposta desse marcador espectral é que ele possa indicar onde a função  $\phi$  atinge valores mínimos dentro de uma vizinhança do sinal.

As componentes de menor magnitude podem corresponder a subharmônicos<sup>1</sup> ou harmônicos menos proeminentes, que são múltiplos fracionários ou menos proeminentes da frequência fundamental e segundo Xiao *et al.* (2019), podem ser relevantes para identificar padrões fora do comportamento esperado. O estudo de componentes espectrais de baixa magnitude é importante para identificar os vales de amplitude nas ondas sonoras. Esses pontos podem representar momentos de menor intensidade na fala, como transições entre palavras ou sílabas menos acentuadas o que pode ser interessante para aplicações de diagnóstico de doenças que afetam a voz. No caso de sinais biológicos como EEG, componentes de baixa magnitude são úteis para identificar períodos de menor atividade elétrica. Esses pontos podem representar momentos de relaxamento ou menor envolvimento cognitivo (YUE *et al.*, 2021).

### 3.2.2.3 O marcador espectral gamma ( $\gamma$ ) proposto

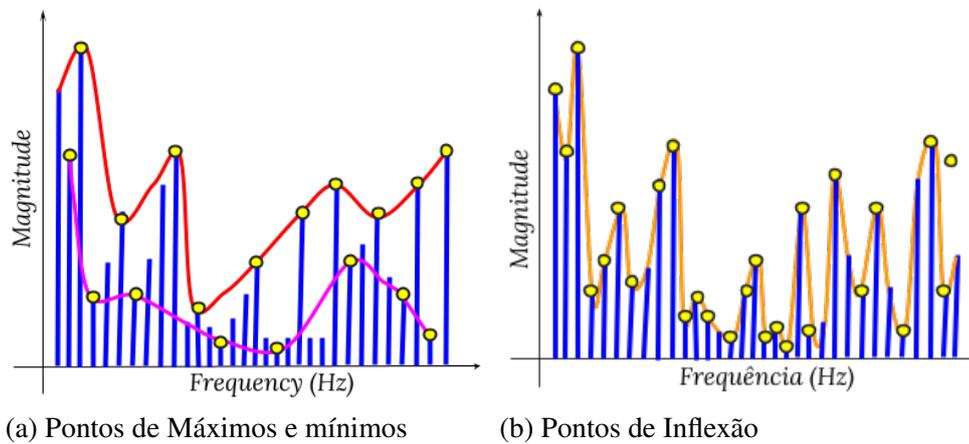
Uma combinação de componentes espectrais de mínimos e máximos também é proposta nesta metodologia. Neste caso, esta metodologia também propõe considerar tanto as altas quanto as baixas magnitudes, o que confere uma maior amostragem de pontos. Este novo conjunto de coeficientes de magnitude define o marcador espectral gamma ( $\gamma$ ) proposto, ilustrados na Figura 12a.

Em ambos os casos, a análise de pontos de máximo e mínimos relativos locais ajudam a identificar eventos de destaque ou características relevantes no sinal. Por exemplo, em sinais de voz, os pontos de máximo podem indicar entonação enquanto ao mesmo tempo magnitude baixas, indicam intervalos na fala ou dificuldade na fala, quando estes eventos acontecem na mesma amostra, a análise de ambos os casos torna-se interessante, como apresenta García *et al.*

<sup>1</sup> Sub-harmônicos são componentes de frequência que aparecem em um sinal periódico e que são submúltiplos inteiros da frequência fundamental. Em outras palavras, um sub-harmônico é uma componente de frequência cuja frequência é menor que a metade da frequência fundamental do sinal.

(2011). Por outro lado, nos exames de EEG, os máximos podem indicar picos de atividade neuronal, tais como estados de concentração e excitação do paciente, associado as magnitudes mínimas, indica momentos de calma mental. Estas transições nos dados, entre máximos e mínimos também contribuem para a compreensão do comportamento desses sinais (CROVATO; SCHUCK, 2007; TATUM, 2021).

Figura 12 – Marcadores espectrais gamma ( $\gamma$ ) e delta ( $\delta$ ) propostos no sinal da função  $\phi$ .



Fonte: Elaborada pelo autor.

#### 3.2.2.4 O marcador espectral delta ( $\delta$ ) proposto

É proposto também a identificação de pontos de inflexão que contribuem na detecção de mudanças na direção da curvatura do sinal. Isso é relevante para entender transições suaves entre diferentes fases do sinal. Em muitos casos, os pontos de inflexão estão associados a eventos específicos nos sinais. Por exemplo, em sinais biomédicos como o eletrocardiograma (ECG), os pontos de inflexão podem estar relacionados a diferentes fases do ciclo cardíaco (YUE *et al.*, 2021). Neste sentido, as mudanças abruptas na curvatura podem indicar tanto comportamentos já caracterizados por tais mudanças quanto também úteis para possíveis anomalias no sinal.

Neste trabalho, é proposto que os pontos de inflexão no sinal  $\phi$  representam as alterações abruptas na magnitude dos componentes espectrais ao longo das frequências. Tais pontos de inflexão consolidam pontos de máximos, mínimos e pontos onde há inversões no crescimento da função  $\phi$ . Esses coeficientes, portanto, representam o conjunto que define o marcador espectral delta ( $\delta$ ) proposto; sua principal diferença em relação aos marcadores de componentes anteriores é que representa mais marcadores à medida que consolidam diferentes magnitudes dos componentes expressos. A Figura 12b destaca os coeficientes do marcador  $\delta$ .

Portanto, embora identificar pontos de inflexão não seja uma terminologia comum à análise no domínio da frequência, a proposta é, de igual modo, identificar as mudanças e transições nas características espectrais, sendo fundamental para detectar transições nas componentes espectrais da curvatura da função  $\phi$ . O que pode indicar transições entre faixas de atenuação e amplificação, comportamentos relevantes para descrever um sinal com padrões espectrais que sofrem mudanças na curvatura da função, tais como ocorre na falha em motores (NAJEEB *et al.*, 2021) e no cenário de detecção da doença de Alzheimer (PARIHAR; SWAMI, 2023). Esse processo de seleção das componentes espectrais através do marcador proposto, detalha, portanto, as etapas 5-7 da Figura 8.

### 3.3 O digrafo como proposta de representação do sinal spectral

Após estabelecer o tipo de marcador nas componentes espectrais do sinal, esta proposta define também como tais marcadores serão processados ao longo dos pontos que formam o sinal no domínio da frequência. Tal análise tem como objetivo final a construção de uma representação estruturada da amostra (e.g. paciente, motor, som), em forma de grafo. Neste sentido, esta seção propõe como as magnitudes espectrais são processadas e interpretadas para a criação do grafo direcionado de arestas ponderadas.

Seja um grafo direcionado  $DG = (V, E)$ , onde  $V = (n_1, n_2, \dots, n_k, \dots, n_L)$ , tal que  $L$  é o número de vértices (ou nós) que estruturam a análise nos marcadores de coeficientes de magnitude, e  $E$  o conjunto de arestas ponderadas tal que dados dois nós,  $n_i$  e  $n_j$ , eles estão conectados no grafo  $DG$  com uma aresta ponderada  $w_{ij} \in E$ , de forma que cada peso  $w_{ij}$  é dado pelo número de vezes que os coeficientes espectrais então contidos em um intervalo de magnitudes  $i$  e apresenta transição para o intervalo de magnitudes  $j$ . Assim, cada transição de mesmo sentido indica que os nós  $n_i$  e  $n_j$  não só estarão ligados por uma aresta direcionada mas também que o número de ocorrências destas transições define o peso  $w_{ij}$  da aresta.

Deste modo, considere  $\Lambda_{min}$  como a amplitude mínima alcançada pelas magnitudes das frequências do sinal  $\phi$ , conforme apresentado na Equação 3.2. E ainda, de acordo com a Equação 3.3, seja  $\Lambda_{max}$  a amplitude máxima alcançada por  $\phi$ , com estas duas definições então é possível descrever a amplitude relativa dos níveis de magnitude spectral a ser representada por cada vértice como  $\Lambda^*$ , de acordo com a Equação 3.4. Notadamente a amplitude relativa indicará a amplitude de cada nível de magnitude a ser representado pelos vértices no Grafo  $DG$ .

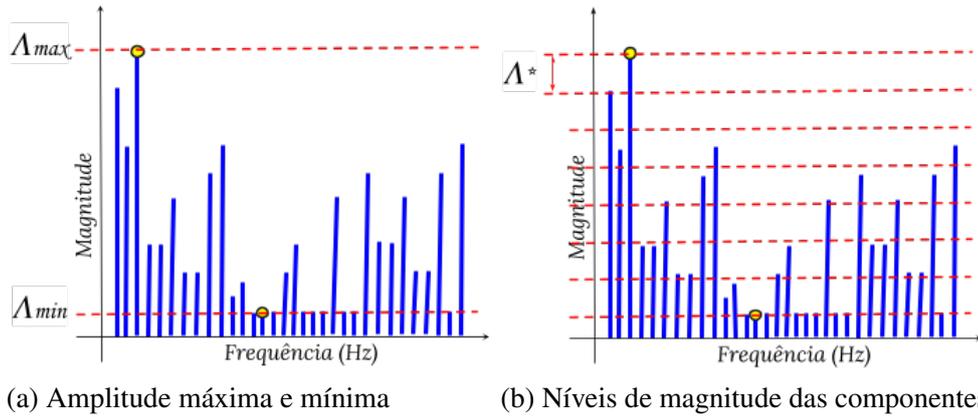
$$\Lambda_{min} = \min\{\phi(k)\} \quad (3.2)$$

$$\Lambda_{max} = \max\{\phi(k)\} \quad (3.3)$$

$$\Lambda^* = \frac{\Lambda_{max} - \Lambda_{min}}{L} \quad (3.4)$$

A Figura 13a ilustra as faixas mínimas e máximas das amplitudes dos coeficientes.

Figura 13 – Amplitude máxima e mínima dos coeficientes e a divisão em níveis de magnitude.



Fonte: Elaborada pelo autor.

O intervalo entre  $\Lambda_{min}$  e  $\Lambda_{max}$  é então dividido em  $L$  intervalos com uma amplitude relativa de  $\Lambda^*$ , cada. Os limites inferior e superior de cada nível  $l$  do intervalo são definidos de acordo com as Equações 3.5 e 6.1.

$$l_{sup}(l) = \Lambda_{min} + l \cdot \Lambda^* \quad (3.5)$$

$$l_{inf}(l) = \Lambda_{min} + (l - 1) \cdot \Lambda^* \quad (3.6)$$

onde  $l \in [1, L]$ .

Deste modo, seja  $range(l)$  uma função que delimita os intervalos de magnitudes no nível  $l$  então dizemos que um marcador, isto é, um coeficiente de magnitude, pertence ao intervalo  $range(l) = [l_{inf}(l), l_{sup}(l)]$  se a  $k$ -ésima componente da DFT, representada por  $\phi(k)$ , apresenta a seguinte propriedade  $P$ :

$$P = \{\phi(k) \in range(l) \mid l_{inf}(l) < \phi(k) \leq l_{sup}(l)\} \quad (3.7)$$

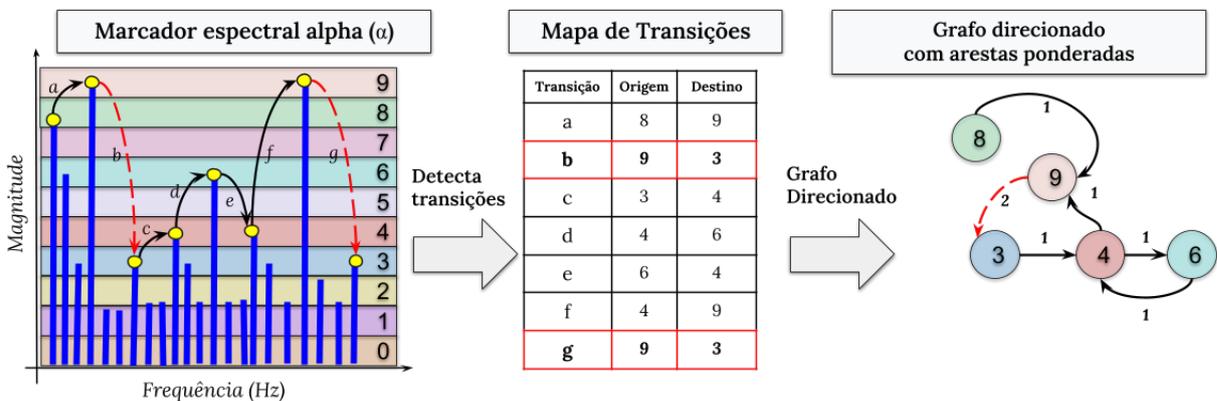
Assim, cada transição de um marcador no nível  $i$  para o nível  $j$  representa uma conexão entre o nó  $n_i$  e o nó  $n_j$ . Vale ressaltar que apenas os coeficientes de magnitude do marcador são considerados para a criação das arestas, observando neste caso o respectivo tipo de marcador espectral adotado dentre àqueles apresentados nas seções anteriores. Assim, após

identificar todas as transições  $t = (n_i, n_j)$  dos marcadores, via mapa de transições, o número de repetições da transição  $t$  define então o valor de  $w_{ij}$ , que representa o peso da aresta  $n_i \rightarrow n_j$ . Esse processo de construção das arestas ponderadas através das transições dos coeficientes espectrais, detalha, portanto, as etapas 8-10 da Figura 8.

Para ilustrar este procedimento, a Figura 14 apresenta a construção do Grafo direcionado com arestas ponderadas, via criação do mapa de transições. Como pode ser observado, temos a construção do mapa de transição referente ao marcador espectral, no exemplo são considerados os coeficientes definidos pelo marcador  $\alpha$ .

A proposta define o mapa de transições com base na ocorrência das componentes espectrais nos diferentes níveis que são representados por vértices. Observe que as transições  $b$  e  $g$  são semelhantes, já que por duas vezes ligam os níveis 9 e 3. Quando isso acontece o peso da aresta deve ser aumentado em mais uma unidade a cada repetição de transição; no exemplo, o peso  $w_{93}$  é então definido como 2, deste modo, dois nós,  $n_9$  e  $n_3$ , então conectados no grafo com a aresta ponderada de peso  $w_{93} = 2$ , conforme o Grafo, à direita. Neste exemplo temos  $L = 10$  e  $k$  de 1 até 10.

Figura 14 – Construção do mapa de transições do marcador no domínio da frequência.



Fonte: Elaborada pelo autor.

### 3.3.1 O processamento multicanal

Por outro lado, esta metodologia também é flexível quanto ao número de canais a serem processados para a extração de atributos. Esta seção apresenta a generalização com um exemplo prático do processamento de vários sinais para representação de um paciente, com base na aquisição de sinais biológicos oriundos de eletrodos em um exame de EEG.

O procedimento descrito na Seção 3.3 é executado para cada sinal. Deste modo, para cada sinal processado temos um digrafo com arestas ponderadas, como apresentado nos passos 1-4 da Figura 15 que ilustra a aplicação da metodologia para o cenário de exames de EEG. Neste caso, destaca-se que temos uma situação que processa múltiplos sinais oriundos de fontes de dados diferentes e que possuem composições de frequências diferentes em seu espectro, contexto conhecido também como multicanal.

Dito isto, como forma de consolidar as estruturas que representam os sinais de cada eletrodo, a metodologia propõe a operação de união, envolvendo cada digrafo, unindo os conjuntos de arestas e vértices, produzindo deste modo, um digrafo que representa, de modo consolidado as características globais, na figura representado pelo grafo  $DG^*$ , passos 5-6 da Figura 15. Seja  $DG_i$  o grafo correspondente ao sinal do eletrodo  $i$  e  $k$  indicando o número de eletrodos, temos então que  $DG^*$  é construído a partir da operação:

$$DG^* = \bigcup_{i=1}^k DG_i \quad (3.8)$$

Com base no grafo  $DG^*$ , são produzidas as matrizes de adjacência (A), que determina quais vértices estão diretamente conectados, e a existência de arestas entre os vértices é fundamental para entender a topologia geral das arestas; além da matriz Laplaciana (L) que representa a relação de conectividade entre os vértices e as arestas direcionadas, destacando componentes fortemente conectados no gráfico DG; e por fim, a matriz Laplaciana Combinatória Direcionada (CL), que apresenta diretamente a estrutura do grafo em termos de conectividade sem enfatizar a direção entre os vértices, passo 7 da Figura 15.

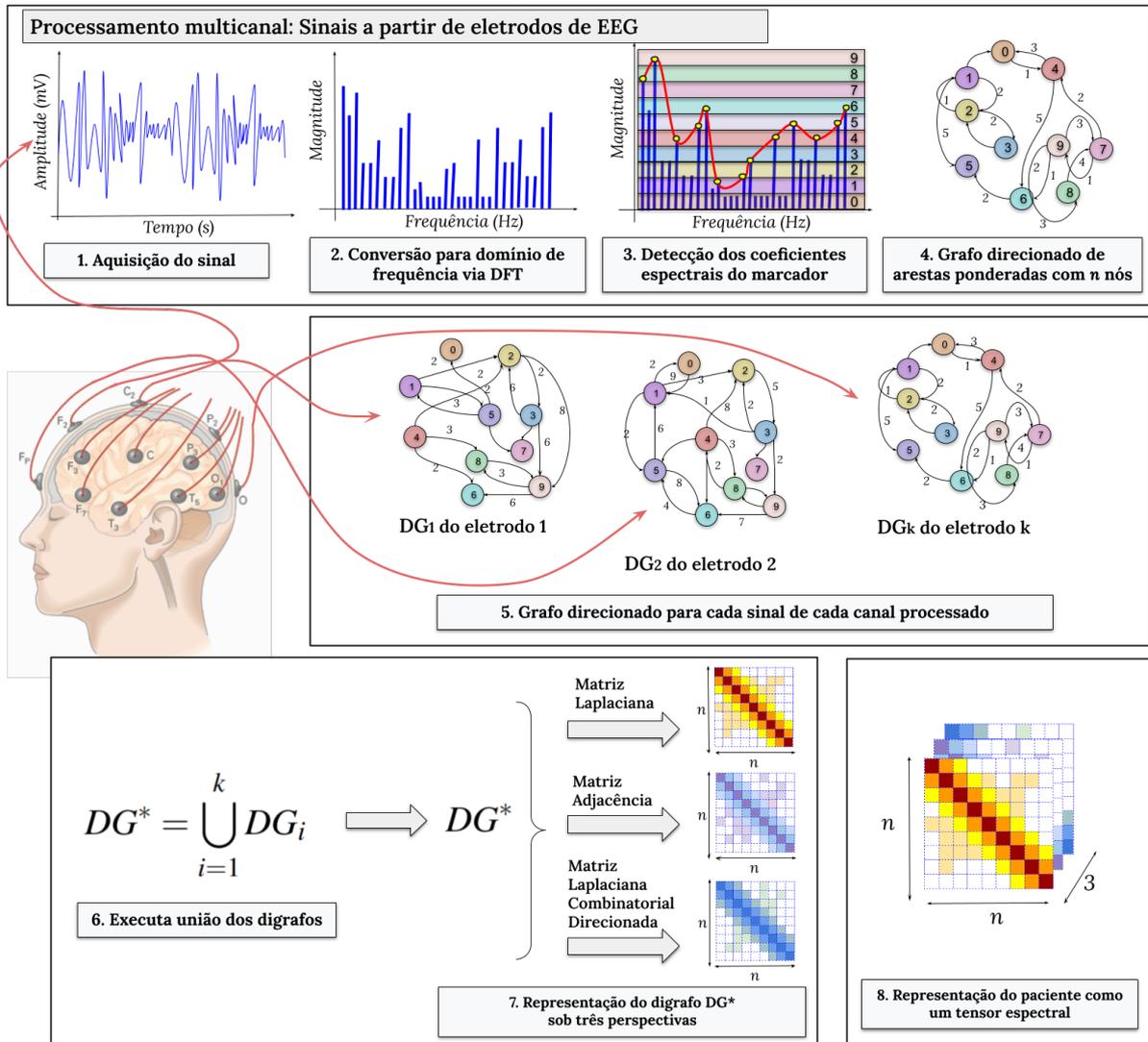
Finalmente, o tensor de três dimensões ( $N \times N \times 3$ ), representando as três matrizes do digrafo com  $N$  vértice, consolida as diferentes perspectivas do digrafo. Cada matriz representa uma perspectiva diferente de análise do digrafo  $DG^*$ . Temos então um único tensor espectral representando a amostra do paciente, Figura 15 (Passos 8).

### 3.3.2 Propriedades da abordagem proposta

Uma vez descrito o procedimento de criação do grafo, revisitamos a definição dos marcadores espectrais  $\alpha$ ,  $\beta$ ,  $\gamma$  e  $\delta$ , de tal modo que possamos exemplificar a criação de diferentes grafos e notar que a escolha do marcador implica não apenas na estrutura do grafo, como na dimensão do número de arestas e vértices, bem como torna mais evidente a análise de características das componentes espectrais.

A Figura 16 apresenta a construção do grafo direcionado  $DG_1$  com base no marcador

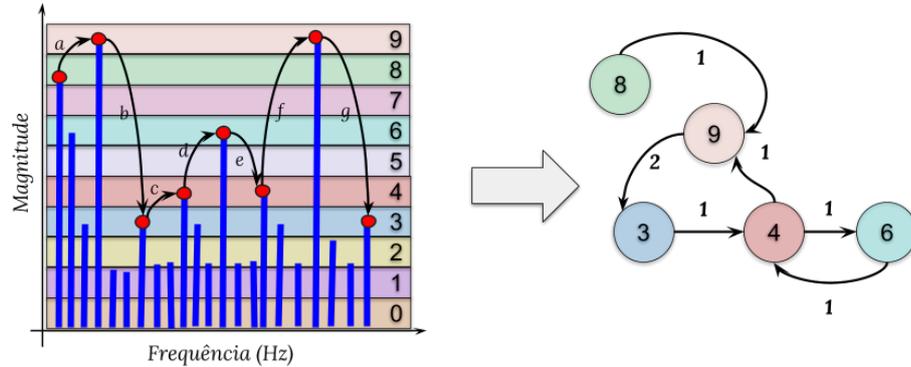
Figura 15 – Fluxo de processamento para cenário multicanal.



Fonte: Elaborada pelo autor.

espectral  $\alpha$ . Este marcador é construído a partir dos coeficientes espectrais nos extremos máximos locais da função  $\phi$ . Em outras palavras, ele representa frequências cujas magnitudes apresentam maiores intensidades. O grafo, à direita, possui um número de vértices  $n(V) = 5$ , com 6 arestas direcionadas.

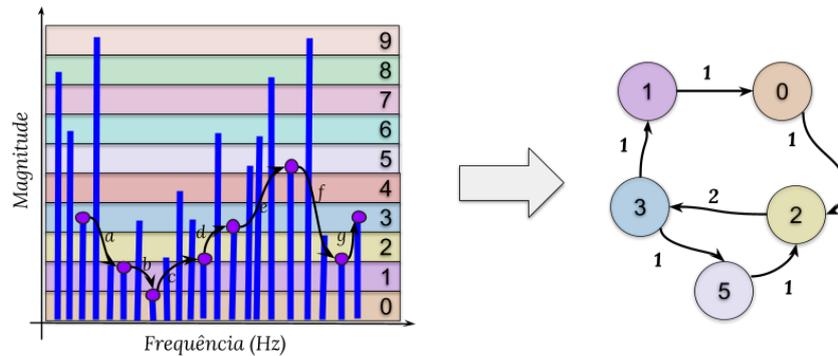
Figura 16 – Grafo direcionado  $DG_1$  construído a partir do marcador  $\alpha$  com  $n(V) = 5$  e  $n(E) = 6$ .



Fonte: Elaborada pelo autor.

Já a Figura 17 apresenta a construção do grafo direcionado  $DG_2$  a partir do marcador espectral  $\beta$ . Tal marcador, por sua vez, indica as frequências de menor intensidade quanto aos seus coeficientes espectrais. Este marcador, diferentemente do marcador  $\alpha$ , denota os extremos mínimos locais da função  $\phi$ . Tais frequências, representam sub-harmônicos que são submúltiplos da frequência fundamental ou harmônicos menos proeminentes e a baixa magnitude pode auxiliar na identificação de variações indesejadas no sinal, tal como ruídos induzidos ou interferências.

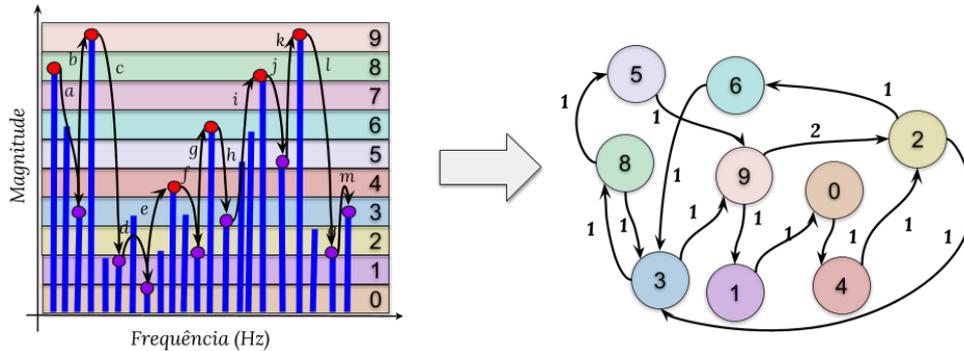
Figura 17 – Grafo direcionado  $DG_2$  construído a partir do marcador  $\beta$  com  $n(V) = 5$  e  $n(E) = 6$ .



Fonte: Elaborada pelo autor.

Por outro lado, a Figura 18 apresenta a construção do grafo direcionado  $DG_3$  a partir do marcador espectral  $\gamma$ . A definição deste marcador considera a união dos coeficientes espectrais representados pelos marcadores  $\alpha$  e  $\beta$ . O grafo construído apresenta uma estrutura maior, com mais arestas e vértices. Essa construção mais complexa se justifica pela maior quantidade de ocorrências de coeficientes espectrais nos diferentes níveis, o que induz a criação de mais arestas entre os vértices, estabelecendo um grafo com ligações mais densas e vértices com graus mais elevados.

Figura 18 – Grafo direcionado  $DG_3$  construído a partir do marcador  $\gamma$  com  $n(V) = 9$  e  $n(E) = 13$ .



Fonte: Elaborada pelo autor.

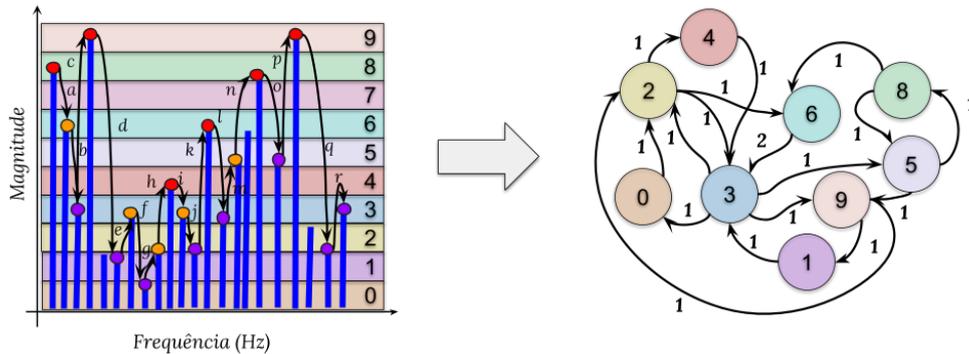
A junção dos marcadores máximo e mínimo apresentada pelo marcador  $\gamma$  visa uma representação do sinal em que ambas as perspectivas estão sendo consideradas. Incorporando neste caso, as vantagens de detectar coeficientes de maior influência sob o ponto de vista de intensidade do sinal original, dado que apresentam magnitudes mais elevadas sem que, ao mesmo tempo, desconsidere as frequências de menor influência no sinal original, tais como frequências que representam pequenas aleatoriedades, anomalias ou interferências.

Também é proposta outra perspectiva de análise das componentes espectrais representada no marcador espectral  $\delta$  proposto. A Figura 19 ilustra o procedimento de análise do sinal e criação do grafo direcionado  $DG_4$ . Nesta representação, é possível observar que tanto os coeficientes espectrais de máximos, em vermelho, quanto os de mínimos locais, em roxo, são considerados, uma vez que estes pontos são extremos máximos locais e também de inflexão. Porém, neste caso, novos coeficientes espectrais são incorporados ao conjunto. Aqueles coeficientes que são identificados nas oscilações de concavidade do sinal em seu domínio da frequência classificam-se como pontos de inflexão, pontos em laranja, e são relevantes para o marcador  $\delta$ . Tais pontos representam magnitudes que não são máximo locais, mas que pela magnitude indicam frequências de magnitude intermediária, como observado no gráfico da figura, à esquerda.

Observar pontos de inflexão no domínio da frequência, embora não seja uma abordagem comum, pode ser útil para indicar transições entre faixas de atenuação e amplificação, comportamentos presentes e úteis de serem identificados em aplicações como falha em motores (NAJEEB *et al.*, 2021) e detecção da doença de Alzheimer (PARIHAR; SWAMI, 2023).

Considerando o grafo da Figura 19, por exemplo, nota-se que o vértice 3 tem maior

Figura 19 – Grafo direcionado  $DG_4$  construído a partir do marcador  $\delta$  com  $n(V) = 9$  e  $n(E) = 16$ .



Fonte: Elaborada pelo autor.

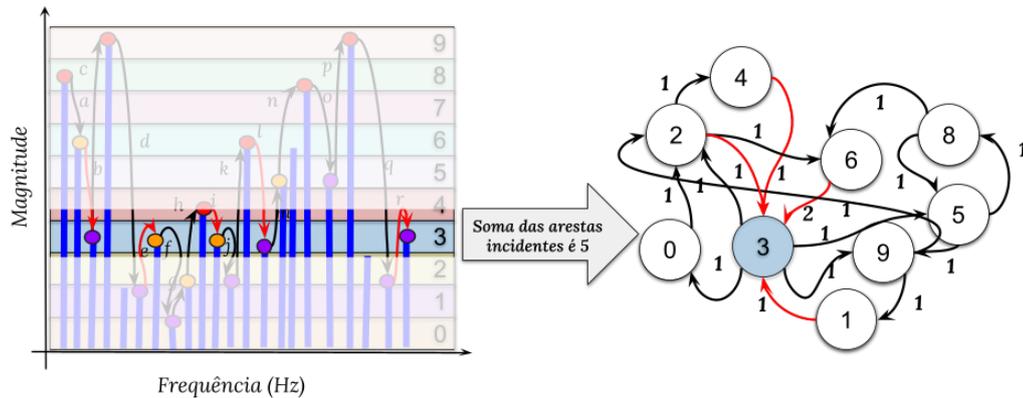
centralidade<sup>2</sup> e neste sentido, ao observar as componentes espectrais do sinal  $\phi$ , à esquerda, nota-se que os marcadores se concentram, de fato, em magnitudes no nível 3. Uma análise proposta é quanto a dispersão das diferentes magnitudes do sinal, o nível do vértice de maior centralidade concentra a maioria dos coeficientes espectrais. Por outro lado, vértices de menor centralidade, podem indicar magnitudes *outliers*.

Como os vértices representam um intervalo de magnitudes, pode-se dizer que a estrutura não é sensível a magnitudes extremas ou a frequências de ruído com magnitude muito baixa, já que a relação entre os coeficientes é representada pelas transições entre níveis, o que indica sua robustez em relação aos extremos. Na estatística descritiva, a média do conjunto é sensível a elementos de valores extremos, por exemplo. A Figura 20, apresenta um caso em que a soma dos pesos das arestas incidentes no vértice 3 denota o nível de magnitude onde há uma maior concentração de coeficientes espectrais, o que também pode ser interpretado como medida de dispersão dos coeficientes. Já o coeficiente de menor magnitude, que poderia representar uma frequência anômala no nível 0 não afeta significativamente a estrutura do grafo.

Além disso, a representação em grafo possibilita também análises topológicas no domínio da frequência, de modo que quando propomos um Passeio sobre as arestas, de modo que um passeio  $P$  partindo do vértice  $n_0$  ao vértice  $n_i$  é uma sequência finita e não vazia de vértices ligados por arestas  $(n_0, n_1, \dots, n_{i-1}, n_i)$ , é possível identificar características relevantes do sinal. Neste sentido, considerando o marcador  $\beta$ , o índice de menor valor em  $P$  indica o nível de magnitude onde então os coeficientes mínimos no sinal  $\phi$ . Esta mesma análise pode ser desenvolvida e aplicada para a identificação de pontos máximos, considerando o marcador  $\alpha$

<sup>2</sup> A centralidade em grafos é uma medida que avalia a importância relativa de um nó ou vértice dentro de um grafo. Existem várias formas de definir medidas de centralidade, dentre elas, a centralidade de grau considera o número de arestas incidentes no vértice, o que indica importância do nó na rede.

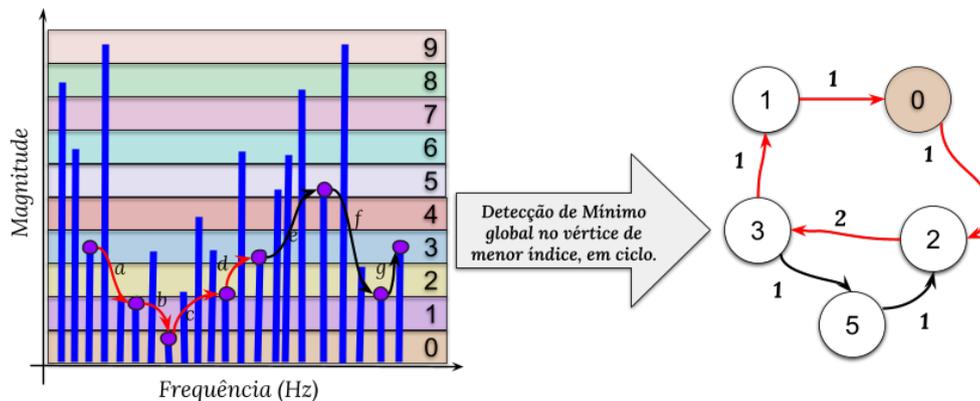
Figura 20 – A soma das arestas incidentes no vértice denota o quantitativo de coeficientes de magnitude presentes no respectivo nível representado.



Fonte: Elaborada pelo autor.

proposto. Para ilustrar, no grafo da Figura 21, com marcadores do tipo  $\beta$ , temos um passeio  $P$  que denota um ciclo entre os vértices 3-1-0-2-3, o vértice de menor índice, neste caso o vértice 0, representa então o intervalo (ou nível) de magnitudes mínimas do sinal  $\phi$ .

Figura 21 – A soma das arestas incidentes no vértice denota o quantitativo de coeficientes de magnitude presentes no nível representado pelo vértice.



Fonte: Elaborada pelo autor.

### 3.3.2.1 A proposta da composição em tensor tridimensional de componentes espectrais

Por outro lado, como apresenta Diestel e Diestel (2017), a representação de grafos por meio de matrizes é uma abordagem fundamental em teoria dos grafos, oferecendo diversas maneiras de capturar as relações e estruturas entre os vértices e arestas. Duas das representações mais comuns são a matriz de adjacência e a matriz de incidência, cada uma com suas próprias vantagens e adequações a diferentes contextos. A matriz de adjacência, por sua vez, apresenta-se como uma matriz quadrada e destaca as relações entre os vértices por meio da presença ou ausência de arestas, conforme descrita na Seção 2.1.3. Existem também outras representações,

como a matriz laplaciana, que reflete propriedades estruturais do grafo. Vale destacar que a escolha entre essas representações depende das características específicas do grafo em questão e das operações que se pretende realizar.

Neste sentido, conforme apresentado nas seções 2.1.3, 2.1.4, 2.1.5 e 2.1.6, as diferentes matrizes são constituídas de características diferentes e por consequência, representam perspectivas diferentes embora do mesmo grafo. A matriz de adjacência ( $A$ ), por exemplo, determina quais vértices estão diretamente conectados, e a existência de arestas entre os vértices é fundamental para a compreensão da topologia geral das conexões do grafo. Já a matriz Laplaciana ( $L$ ) representa a relação de conectividade entre os vértices e as arestas direcionadas, destacando componentes fortemente conectados no grafo DG. Por fim, a matriz Laplaciana Combinatória Direcionada ( $LD$ ) apresenta diretamente a estrutura do grafo em termos de conectividade sem enfatizar a direção entre os vértices.

Neste sentido, ao representar a amostra do sinal  $\phi$  como tensor podemos combinar as diferentes perspectivas do grafo, representadas por cada matriz individualmente. Assim, a estrutura em tensor é proposta como mecanismo estruturado para finalmente representar o sinal original de maneira consolidada. Sejam  $A$ ,  $L$ , e  $LD$  matrizes  $n \times n$  representando a perspectiva do grafo como matriz de adjacência, como matriz laplaciana, e como matriz laplaciana combinatória direcionada, respectivamente, temos:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \quad (3.9)$$

$$L = \begin{bmatrix} l_{11} & l_{12} & \dots & l_{1n} \\ l_{21} & l_{22} & \dots & l_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \dots & l_{nn} \end{bmatrix} \quad (3.10)$$

$$LD = \begin{bmatrix} ld_{11} & ld_{12} & \dots & ld_{1n} \\ ld_{21} & ld_{22} & \dots & ld_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ld_{n1} & ld_{n2} & \dots & ld_{nn} \end{bmatrix} \quad (3.11)$$

A operação de concatenação ao longo de uma nova dimensão cria um tensor tridimensional  $T$  de dimensões  $n \times n \times 3$ :

$$T = \begin{bmatrix} A \\ L \\ LD \end{bmatrix} \quad (3.12)$$

Cada bloco de  $n \times n$  elementos representa uma das matrizes e conseqüentemente preserva em seu bloco as informações de sua própria perspectiva sobre o grafo, já as colunas e linhas dentro de cada bloco correspondem aos vértices no grafo bem como os valores representam suas relações. Esse processo de representação das matrizes consolidadas no tensor espectral que finalmente representa a amostra do paciente, detalha, portanto, as etapas 11-13 da Figura 8.

Diante deste cenário, representar uma amostra como um tensor tridimensional  $n \times n \times 3$  oferece várias vantagens do ponto de vista computacional, sendo particularmente útil em operações de reconhecimento de padrões, especialmente em tarefas de classificação de amostras no contexto de processamento em redes neurais profundas, o que possibilita também agregar vantagens como treinamentos de modelos usando transferência de aprendizagem ao utilizar modelos pré-treinados com outros tensores.

Dentre as vantagens da representação proposta, também destaca-se que a representação tridimensional permite analisar a estrutura topológica do grafo em cada uma das três dimensões, tal como ocorre nos diferentes canais de cores no contexto de classificação de imagens. Isso é crucial para reconhecimento de padrões, pois as diferentes perspectivas contida em cada camada do tensor, tais como aspectos visuais em diferentes cores, possibilitam analisar de diferentes formas os vértices das componentes conexas do grafo e suas relações.

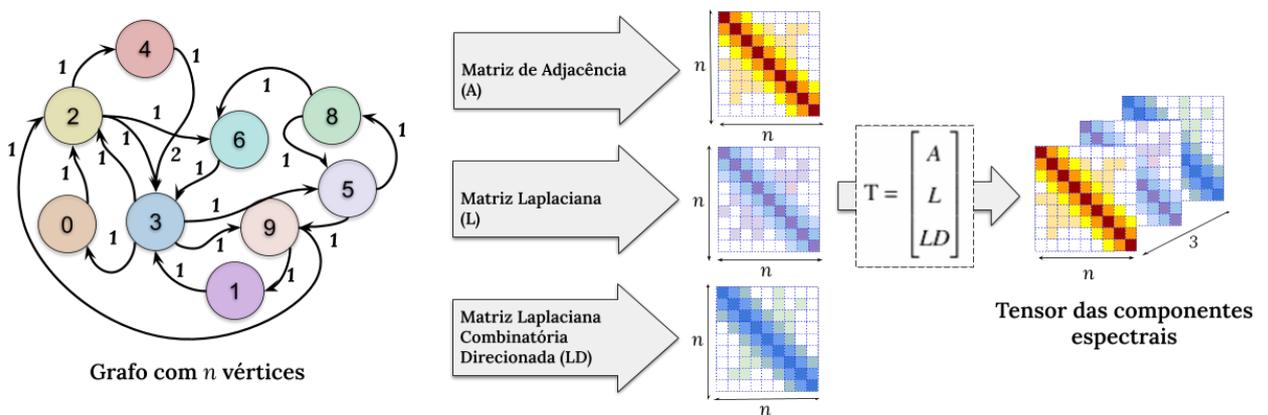
Outra vantagem da representação tensorial é sua compatibilidade com operações convolutivas, uma vez que tensores tridimensionais são bem adaptados para operações de convolução envolvendo os múltiplos neurônios nas diferentes camadas das CNNs. Como a relação entre os vértices nas matrizes é preservada, as convoluções podem detectar padrões locais em diferentes partes do grafo e a representação  $n \times n \times 3$  permite que essas operações sejam aplicadas eficientemente em cada matriz.

Por outro lado, destaca-se que a facilidade de manipulação e processamento que o uso de tensores proporciona para a manipulação e processamento de dados, especialmente com bibliotecas de aprendizado de máquina como TensorFlow e PyTorch. Essas bibliotecas são

otimizadas para trabalhar com tensores, fomentando a implementação de novos algoritmos de extração e modelos de classificação.

Em resumo, a representação de uma amostra como um tensor de dimensão  $n \times n \times 3$  oferece uma representação rica e estruturada, aproveitando as características espaciais, preservando as ligações entre vértices, bem como os valores das arestas do grafo. A Figura 22 ilustra a etapa descrita na construção do tensor espectral. Cada uma das três matrizes são consolidadas em uma única estrutura tensorial denominada tensor espectral:

Figura 22 – Junção das diferentes perspectivas do grafo representadas por cada matriz em uma estrutura tensorial denominada tensor espectral.



Fonte: Elaborada pelo autor.

Após os marcadores dos coeficientes de magnitude identificados, passando pela criação do mapa de transição e cálculo do peso das arestas, o grafo direcionado  $DG$  é construído considerando os diferentes níveis de amplitude do sinal. O grafo  $DG$ , que originalmente era representado individualmente pelas matrizes de Adjacência, Matriz Laplaciana e Matriz Laplaciana Combinatória, agora está consolidado em um único tensor espectral, constituindo portanto a estrutura final do descritor da amostra.

### 3.4 Algoritmo do descritor de sinais proposto baseado em grafo direcionado

Uma vez descrita as etapas fundamentais da metodologia, nesta seção será apresentado o pseudocódigo do algoritmo descritor do sinal, para a criação do grafo direcionado e por conseguinte, o tensor espectral. O algoritmo 1 apresenta o pseudocódigo para a construção do grafo direcionado. Inicialmente, com base no sinal e frequência recebidos como parâmetros, é calculada a DFT para identificar as frequências e coeficientes de magnitude do sinal, produzindo

a versão no domínio da frequência, *senal\_phi*. De posse do *senal\_phi* são calculados os limites máximo e mínimos de magnitudes, além de calculada a amplitude a ser usada em cada nível, tais níveis serão apresentados por cada vértice (Linhas 2 - 6).

O passo seguinte é definir quais intervalos de magnitude serão representados pelos vértice do grafo. A variável *range\_vertice* é uma lista de vetores que definem o limite inferior e superior representado por cada vértice (Linhas 7-12). Uma vez definidos os intervalos, o algoritmo avança para a etapa fundamental que é a identificação dos coeficientes espectrais que representam os marcadores do sinal. Conforme o tipo de marcador, a função *extrai\_marcadores* executa o procedimento descrito na seção 3.2 para identificar as componentes espectrais do sinal, ao término, a variável *marcadores\_phi* armazena as posições no sinal *senal\_phi* onde foram identificadas os coeficientes espectrais que representam os marcadores (Linha 16).

Finalmente, a construção do grafo se dá pela definição do conjunto de arestas ponderadas  $E$ . Este conjunto armazena tuplas  $(u, v, w_{uv})$  que definem as ligações entre os vértices  $u$  e  $v$  com aresta de peso  $w_{uv}$ . A construção das arestas acontece percorrendo toda a lista de marcadores encontrados no passo anterior, e com base na função *get\_vertice*, que identifica o vértice onde houve a ocorrência do marcador, a variável de controle *vertice\_v* é definida.

A variável *vertice\_v* representa a primeira ligação entre vértices caso o vértice  $u$  ainda não exista, caso contrário, o vértice  $v$  é então ligado ao vértice  $u$ . Se a aresta  $u \rightarrow v$  não existir no conjunto  $E$ , então ela é criada com peso  $w_{uv} = 1$ , caso ela já exista, temos então um caso onde há repetição de transição entre níveis, situação descrita na seção 3.3, neste caso o peso da aresta é incrementado em uma unidade. Ao término, o algoritmo finaliza com o conjunto  $E$  construído. Temos os vértices, as arestas e seus respectivos pesos que definem o grafo.

---

**Algorithm 1** Algoritmo para a construção do grafo direcionado de arestas ponderadas
 

---

```

1: Entrada: sinal, frequencia, num_vertices, marcador
   /* Calcula DFT e descarta os números complexos */
2: sinal_phi ← abs(DFT(sinal,frequencia))
   /* Identifica magnitude máxima e mínima do sinal phi */
3: sinal_phi_max ← max(sinal_phi)
4: sinal_phi_min ← min(sinal_phi)
5: sinal_phi_ampl ←  $(sinal\_phi\_max - sinal\_phi\_min) / num\_vertices$ 
6: range_vertices ← [ ]
   /* Calcula os intervalos, inferior e superior, de magnitude que serão
   representados por cada vértice. */
7: for l ← 1 to nodes_max do
8:
9:   sinal_phi_lsup ← sinal_phi_min + l * sinal_phi_ampl
10:  sinal_phi_linf ← sinal_phi_min + (l - 1) * sinal_phi_ampl
11:  range_vertices[l] ← [sinal_phi_linf, sinal_phi_lsup]
12: end for
   /* Inicializa os vértices de controle e o conjunto de arestas E */
13: vertice_u ← Null
14: vertice_v ← Null
15: E ← Null
   /* Identifica os pontos de marcadores espectrais, nas frequências do sinal  $\phi$ 
   conforme o tipo de marcador. */
16: marcadores_phi ← extrai_marcadores(sinal_phi, marcador)
   /* Calcula o peso das arestas conforme a ocorrência de cada marcador no range de
   cada vertice. */
17: for k ← 0 to marcadores_phi.length - 1 do
18:
19:   /* Retorna o vértice que corresponde ao intervalo onde está o marcador. */
   vertice_v ← get_vertice(sinal_phi[marcadores_phi[k]], range_vertices)

20:   if vertice_u = Null then
   /* Se vértice u não existir. */
   vertice_u ← vertice_v

21:22: else
23:   if (vertice_u, vertice_v) not in E then
   /* Inicializa a aresta u→v com peso  $w_{uv} = 1$  */
   adiciona_aresta(E, (vertice_u, vertice_v), 1)

24:25: else
   /* Atualiza o peso da aresta u→v com peso  $w_{uv} = w_{uv} + 1$  */
   atualiza_aresta(E, (vertice_u, vertice_v), 1)

26:27: end if
28:   vertice_u ← vertice_v
29: end if
30: end for

```

---

### 3.4.1 Complexidade computacional do algoritmo proposto

Além da construção, é pertinente considerar os custos computacionais do algoritmo proposto. Neste sentido, a análise de algoritmo também propõe o cálculo do custo de algoritmos com base na notação assintótica que leva em consideração situações extremas, isto é, quando o número de amostras com as quais ele tem que atuar é consideravelmente grande. Neste sentido, considerando como custo constante as estruturas que estão associados a operações mais simples, tais como atribuições, cálculos aritméticos e checagem de condicionais além de estruturas de repetições as quais o custo computacional não aumenta conforme o número de amostras.

O algoritmo proposto para a extração de atributos do sinal, inicia aplicando a DFT que processa com base em um sinal com  $n$  amostras (Linha 2). O custo computacional da DFT é comumente associado a ordem  $O(n^2)$ . Entretanto, segundo Bekele (2016), é possível desempenhar o processamento da DFT a um custo computacional inferior, na ordem de  $O(n \log(n))$ . Os passos que seguem, representam operações simples para busca de máximos e mínimos cujo custo depende linearmente do número de amostras no sinal, portanto, custo na ordem de  $2 \cdot O(n)$ . Avançando, já na etapa de cálculo dos intervalos (Linhas 7-12) temos uma estrutura de repetição que não depende do número de amostras, na realidade ela apresenta um custo constante proporcional ao número de vértices, que não cresce assintoticamente conforme o tamanho da amostra, portanto consideremos como custo constante  $O(1)$ .

Já a etapa de detecção de marcadores (Linha 16), é necessário percorrer todo o sinal, processando cada uma das  $n$  amostras do sinal *senal\_phi*. Novamente, considerando o pior caso, temos um custo proporcional a  $O(n)$ . Por fim, a etapa de criação das arestas ponderadas executa em um número de vezes proporcional a quantidade de marcadores encontrados. Se considerarmos o pior caso em que este número é no máximo igual ao total de amostras, novamente teremos o custo linear de  $O(n)$ . Finalmente, considerando todos os custos individuais de cada trecho do algoritmo de extração temos o custo total  $T$  que é assintoticamente proporcional a:

$$\begin{aligned}
 T &= O(n \log n) + 2 \cdot O(n) + O(1) + O(n) + O(n) \\
 T &= O(n \log n) + 4 \cdot O(n) + O(1) \\
 T &= O(n \log n)
 \end{aligned} \tag{3.13}$$

Concluindo, portanto, que o algoritmo possui um custo assintótico na ordem de  $O(n \log n)$ , onde  $n$  refere-se ao número de amostras do sinal de entrada.

Já para a fase de composição em tensores, o custo computacional para calcular a matriz Laplaciana de um grafo direcionado depende do número de vértices ( $K$ ) e da quantidade de arestas presentes no grafo ( $|E|$ ). A matriz Laplaciana, por exemplo, dependerá do cálculo da matriz de adjacência do grafo direcionado, uma vez que ela é obtida subtraindo a matriz de adjacência pela matriz de grau.

Sendo  $K$  a ordem do grafo, o custo computacional para calcular a matriz Laplaciana pode ser aproximadamente  $O(K^2 + |E|)$ . A parte  $O(K^2)$  está relacionada à construção da matriz de adjacência e à matriz de grau, enquanto a parte  $O(|E|)$  é devido à contagem das arestas, este custo deve ser duplicado, já que consideramos também a matriz Laplaciana Combinatória Direcionada. Finalmente, adicionando este custo a Equação 3.13, temos o custo total proporcional a:

$$T = O(n \log n) + 2O(K^2 + |E|) \quad (3.14)$$

onde  $n$  é o número de amostras no sinal,  $K$  é o número máximo de vértices do grafo direcionado e  $|E|$  denota o número de arestas. Embora quadrático, é importante notar que  $K$  não cresce na mesma ordem que o número  $n$  do tamanho da entrada do algoritmo.

### 3.5 Algoritmo proposto para seleção dos parâmetros

Por sua vez, o algoritmo 2 apresenta o passo a passo para a seleção dos parâmetros do método de extração. O algoritmo recebe uma partição do conjunto de dados no parâmetro chamado *dataset*, este conjunto será utilizado para encontrar a melhor combinação dos dois parâmetros do método: número de vértices e tipo de marcador.

Além do parâmetro *dataset* o método também recebe o *max\_vertices* que estabelece o limite superior para as combinações com número de vértices que serão testadas; além do parâmetro *classificador* que neste caso representa um modelo capaz de processar a estrutura de tensor para a tarefa de classificação. A proposta do algoritmo de seleção é baseada na busca em grade, variando os valores de ambos os parâmetros no sentido de obter um melhor *score* médio, a partir das taxas de acerto. Esta métrica é obtida por meio do mecanismo de validação cruzada estratificada<sup>3</sup>, que é particularmente útil em cenários onde pode existir desbalanceamento entre as classes.

<sup>3</sup> A validação cruzada estratificada ajuda a mitigar o viés que pode ocorrer ao avaliar o desempenho de um modelo em um único conjunto de dados. Nesse método as proporções das classes nos folds são mantidas, garantindo que cada fold tenha uma distribuição semelhante das classes ao conjunto de dados original.

---

**Algorithm 2** Algoritmo para a seleção do número de vértices do extrator proposto
 

---

```

1: Input: dataset, max_vertices, passo = 3, classificador
2: ultimo_score ← 0
3: min_vertices ← 3
4: n_splits ← 3
5: melhor_num_vertices ← min_vertices
   /* Vetor com os marcadores possíveis. */
6: marcadores ← [ $\alpha, \beta, \gamma, \delta$ ]
   /* Divisão do dataset para validação cruzada estratificada. */
7: kfold, x, y ← StratifiedKFold(n_splits, dataset)
   /* Executa uma busca em grade pelo número de vértices. */
8: for num_vertices ← min_vertices to max_vertices do
9:   /* Executa uma busca em grade pelo marcador de melhor score. */
10:  for m ← 0 to marcadores.length - 1 do
11:    /* Executa a extração de características no sinal e consolida as matrizes no
       tensor, para cada amostra em x. */
12:    x_tensor ← extractSpectralTensor(x, num_vertices, marcadores[m])
       /* Executa validação cruzada estratificada ao numero de n_splits vezes. */
13:    estatisticas ← crossValidation(classificador, x_tensor, y, kFold)
14:    media_score ← media(estatisticas.acuracia)

15:    if media_score ≥ ultimo_score then
16:      ultimo_score ← media_score
17:      melhor_num_vertices ← num_vertices
18:    end if
19:  end for
20:  num_vertices ← num_vertices + passo
21: end for

```

---

### 3.5.1 Complexidade computacional do algoritmo proposto

A linhas iniciais do algoritmo apresentam operações básicas com custo computacional constante,  $O(1)$ . Entretanto, já na linha 7, temos a divisão do dataset em partições para treino e teste do modelo. Esta etapa analisa todas as amostras contidas no dataset. Considerando  $n$  como o total de amostras, o custo deste método então é proporcional a  $O(n)$ . Mais adiante, no trecho seguinte têm-se duas estruturas de repetições aninhadas. A primeira apresenta uma variação proporcional  $K$  máximo de vértices, portanto proporcional a  $O(K)$ . Já o laço interno varia com o número  $m$  de marcadores que neste caso é constante,  $O(m)$ . As linhas de 12 a 13 executam ambas funções cujo custo é proporcional ao número de amostras  $n$ , temos um custo de  $O(2n + c)$ , onde  $c$  é o custo das operações de valor constante (Linhas 15-18). Consolidando

todos os trechos, temos então o custo  $T$ :

$$\begin{aligned} T &= O(n) + O(k) \cdot O(m) \cdot O(2n + c) \\ T &= O(n) + O(2n + c) \\ T &= O(n) \end{aligned} \tag{3.15}$$

### 3.6 Conclusão do Capítulo

No encerramento deste capítulo, apresentamos a metodologia proposta nesta tese como um estratégia que inova no contexto da extração de características de sinais no domínio da frequência, com ênfase na análise espectral, desenvolve uma estrutura de grafo direcionado, também conhecido como digrafo, que é representada em diferentes aspectos consolidados em um tensorial espectral do sinal ou do conjunto de sinais processados. A abordagem proposta se destaca pelo digrafo cujas arestas são ponderadas conforme os coeficientes espectrais processados e refletem as características do sinal.

A proposta inicia na transformação do sinal via DFT, produzindo um ambiente para o processamento no domínio da frequência, o sinal de entrada é processado dentro de uma perspectiva que o quantiza em  $L$  níveis representados pelos vértices do digrafo. A proposta também sugere marcadores das componentes espectrais sob quatro perspectivas distintas: pontos de máximos locais, propostos como marcador espectral  $\alpha$ ; pontos de mínimos locais, propostos como marcador espectral  $\beta$ , além da união dos pontos de máximos e mínimos, propostos como marcador espectral  $\gamma$ ; e, por último, pontos de inflexão, propostos como marcador espectral  $\delta$ .

O grafo construído é então representado por meio de três matrizes essenciais: adjacência, laplaciana e laplaciana combinatória direcionada. A combinação dessas matrizes resulta em um tensor único que encapsula as informações do sinal. A proposta é que este tensor possa ser eficazmente empregado como uma representação robusta da amostra, podendo ser combinado com classificadores que operam com tensores como entrada. A integração desses elementos visa não apenas aprimorar a representação do sinal, mas também potencializar a eficácia de métodos de classificação que se beneficiam de informações multidimensionais e estruturadas em tensores. Este avanço metodológico abre novas perspectivas na análise de sinais no domínio da frequência, contribuindo para a inovação e discernimento na interpretação de padrões espectrais complexos.

## 4 RESULTADOS NO CONTEXTO DE AMOSTRAS SINTÉTICAS

Neste capítulo, apresentamos a metodologia aplicada a amostras de sinais produzidos de maneira sintética. A proposta é apresentar a criação digrafo sobre diferentes sinais, além de ilustrar as principais diferenças do grafo descritor sob o domínio do tempo e sob o domínio da frequência. Ao final, apresentamos o comparativo entre duas amostras de classes diferentes e comparamos a estrutura dos grafos produzidos em diferentes cenários.

### 4.1 Amostras sintéticas

A partir das definições matemáticas da função seno e a possibilidade de compor um sinal com base na junção de diferentes frequências, característica esta que representa um cenário mais próximo ao comum do estado de um sinal real, é possível então expressar matematicamente a construção do sinal, incluindo o trecho da inclusão do ruído aleatório para que haja diferenciação entre as amostras da mesma classe. Deste modo, considerando  $x(t)$  a função que representa o sinal senoidal, podemos expressá-la como:

$$x(t) = \sum_{i=1}^N A_i \cdot \sin(2\pi f_i t) \quad (4.1)$$

onde  $x(t)$  é o sinal gerado no tempo,  $N$  é o número de componentes senoidais,  $A_i$  é a amplitude da  $i$ -ésima componente senoidal,  $f_i$  é a frequência da  $i$ -ésima componente senoidal, e por fim,  $t$  é a variável de tempo.

Assim, visando representar variações aleatórias de aquisição e ambiente, a função  $x(t)$  com a inclusão do ruído aleatório pode ser expressa como:

$$x_{\text{noisy}}(t) = x(t) + n(t) \quad (4.2)$$

onde  $x_{\text{noisy}}(t)$  é o sinal com ruído,  $x(t)$  é o sinal sem ruído e  $n(t)$  é o ruído aleatório.

Neste caso, a representação do ruído  $n(t)$  é modelado como uma variável aleatória gaussiana com média zero e desvio padrão  $\sigma$ , isto é,  $n(t) \sim \mathcal{N}(0, \sigma^2)$ . Finalmente, a expressão completa para a geração do sinal com ruído é:

$$x_{\text{noisy}}(t) = \sum_{i=1}^N A_i \cdot \sin(2\pi f_i t) + n(t) \quad (4.3)$$

onde  $x_{\text{noisy}}(t)$  é o sinal gerado com ruído,  $A_i$  e  $f_i$  são as amplitudes e frequências das componentes senoidais, respectivamente, já  $n(t)$  é o ruído gaussiano.

A partir das definições anteriores, criou-se então dois conjuntos distintos de sinais. Com base na Equação 4.3, o primeiro conjunto é construído com as amplitudes  $A_{\text{classe0}}$  e

frequências  $f_{classe0}$ , descritas por:

$$A_{classe0} = [1, 0.8, 0.6, 0.4, 0.3] \quad (4.4)$$

$$f_{classe0} = [5, 50, 150, 300, 400] \quad (4.5)$$

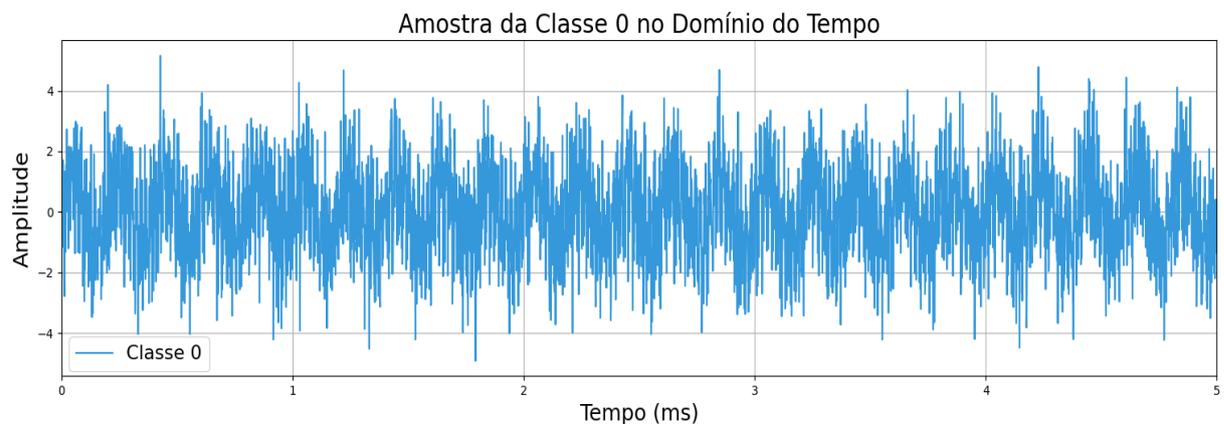
Já o segundo conjunto, é definido pelas amplitudes  $A_{classe1}$  e frequências  $f_{classe1}$ , descritas por:

$$A_{classe1} = [1, 0.8, 0.6, 0.4, 0.3, 0.2, 0.15, 0.1, 0.08, 0.05] \quad (4.6)$$

$$f_{classe1} = [10, 30, 70, 100, 150, 200, 250, 350, 450, 600] \quad (4.7)$$

A Figuras 23 e 24 ilustram um exemplo de um sinal de cada classe. Na cor azul, é apresentada a amostra formada por 5 diferentes frequências do conjunto  $f_{classe0}$ , já na cor vermelha uma amostra do sinal formado por 10 frequências, definido pelo conjunto  $f_{classe1}$ . Em ambos os casos, cada frequência com suas respectivas magnitudes.

Figura 23 – Amostra do sinal sintético formado por 5 frequências com magnitudes distintas.

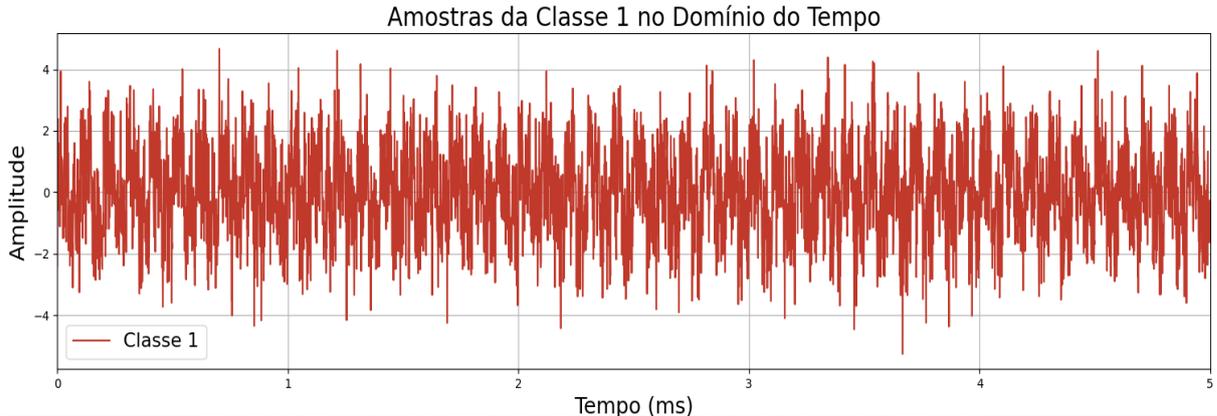


Fonte: Elaborada pelo autor.

Na figuras anteriores, os sinais resultantes exibem amplitudes e frequências distintas, embora mostrem semelhança visual no domínio do tempo. Essa similaridade visual pode estar refletida também na extração de características e prejudicar o desempenho de modelos de classificação. Para superar esse desafio, a transformação do sinal via DFT proporciona uma perspectiva alternativa, destacando as relações de frequências e magnitudes dos sinais, como ilustrado na Figura 25, apresentando os mesmos sinais no domínio da frequência.

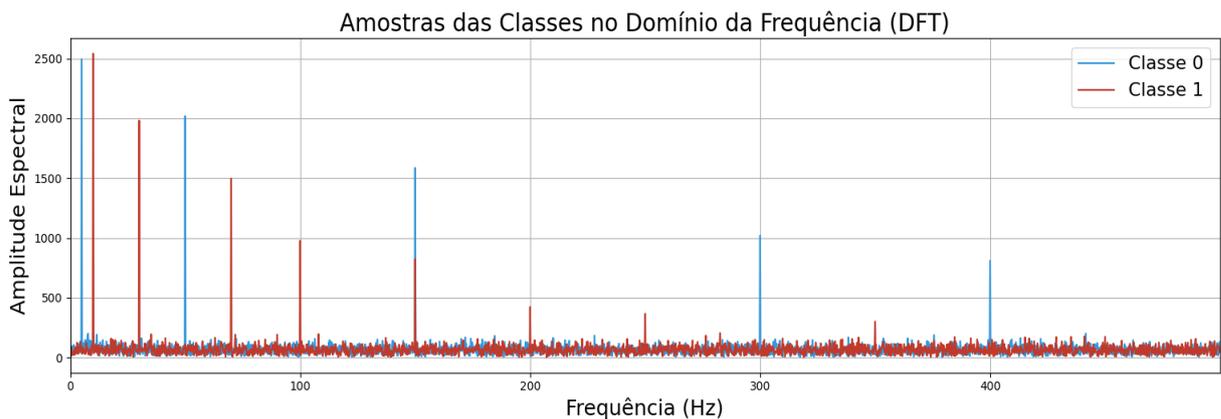
Como apresentado no Algoritmo 1, após o sinal ser convertido para o domínio da frequência, a próxima fase é estabelecer o número de níveis de magnitude do sinal que serão

Figura 24 – Amostra do sinal sintético formado por 10 frequências com magnitudes distintas.



Fonte: Elaborada pelo autor.

Figura 25 – Amostras sintéticas no domínio da frequência, em azul o sinal formado por 5 frequências, já em vermelho o sinal formado por 10 frequências distintas.



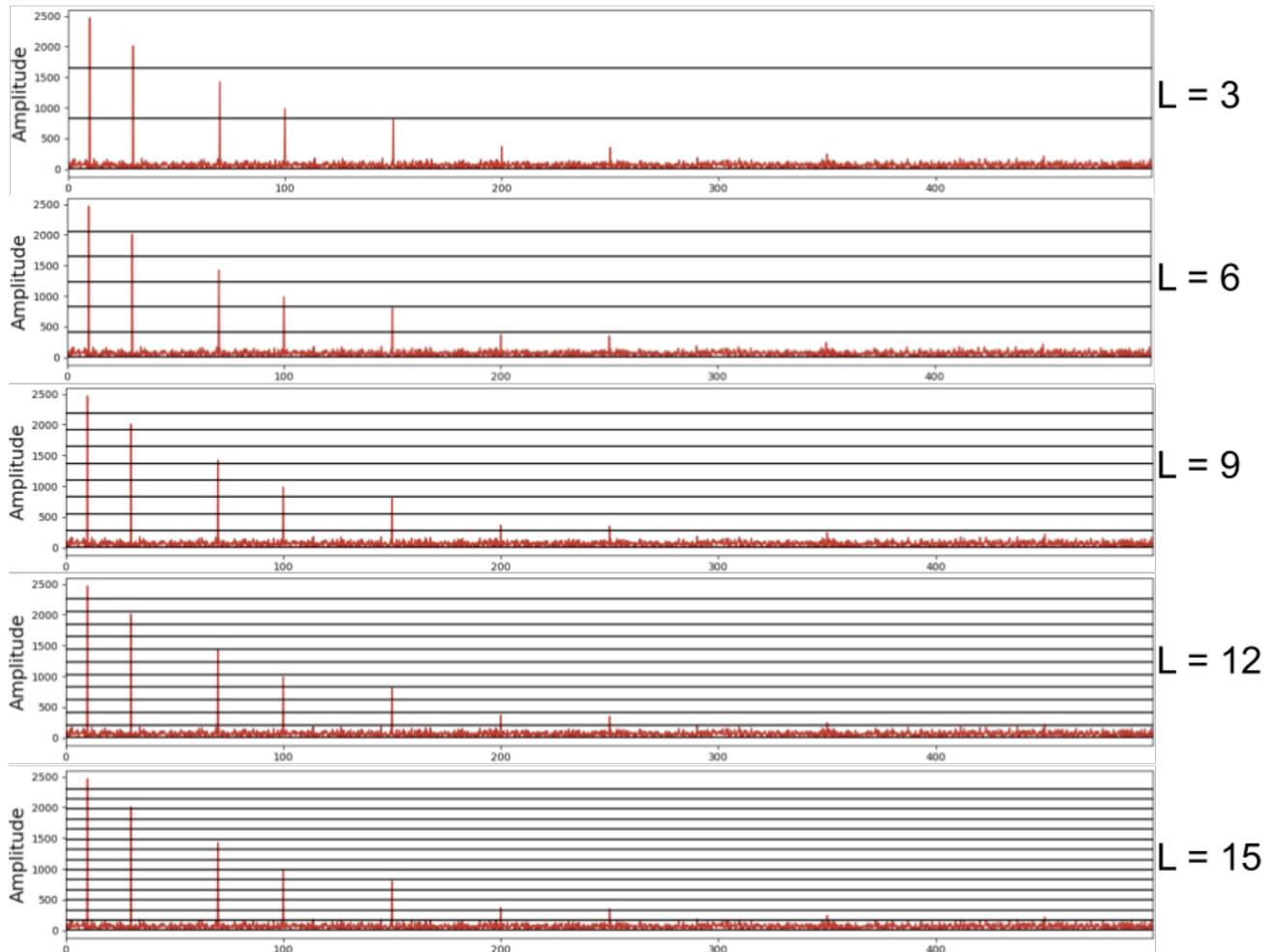
Fonte: Elaborada pelo autor.

representados pelos vértices do grafo. Neste sentido, para fins comparativos, a Figura 26 ilustra a separação em níveis, considerando  $L$  iniciando com 3 e variando até 15, com passo de 3.

## 4.2 Construção do digrafo como descritor do sinal

Avaliando as diferentes perspectivas de domínio do mesmo sinal, a Figura 27 apresenta um comparativo entre digrafos criados sobre o sinal no domínio do tempo em comparação com o domínio da frequência. Neste sentido, é possível observar que mesmo considerando a quantidade de níveis  $L$  similar, é notória a diferença de complexidade estrutural entre os digrafos. O digrafo da Figura 27a, por exemplo, apresenta 20 arestas distribuídas entre 6 vértices, em contrapartida, temos no mesmo nível  $L$  apenas 12 arestas e 5 vértices para a representação do

Figura 26 – Exemplo da divisão em níveis de magnitude do sinal sintético da classe 1, variando  $L$  de 3 até 15. Cada nível é representando por um vértices no digrafo.



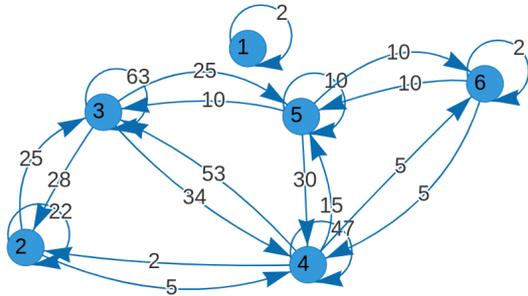
Fonte: Elaborada pelo autor.

mesmo sinal no domínio da frequência. Já a Figura 27c, com  $L = 12$ , apresenta um digrafo com 45 arestas e 12 vértices, para o domínio do tempo, enquanto o mesmo sinal é representado por 9 arestas e 5 vértices, no domínio da frequência, Figura 27d.

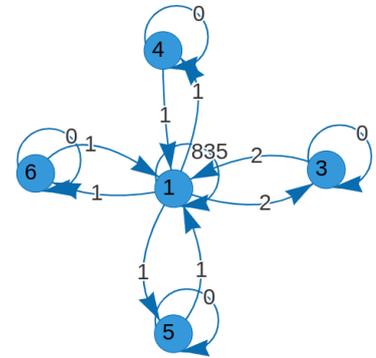
Além disso, considerando o marcador espectral  $\alpha$ , que toma como base os extremos máximos locais do sinal, é possível construir o digrafo baseado nas componentes espectrais. A Figura 28-a, apresenta, em azul, o digrafo correspondente a um sinal da classe 0, considerando um  $L = 3$ . Por outro lado, a Figura 28-b, apresentam também um grafo que representa uma amostra da classe 1. Nota-se uma ligeira diferença nos pesos das arestas.

Quanto a Figura 28, é possível observar que os digrafos construídos apresentam características interessantes. Destacamos, por exemplo, a possibilidade de melhor identificar o intervalo de magnitudes com maior concentração de coeficientes, que neste caso representam

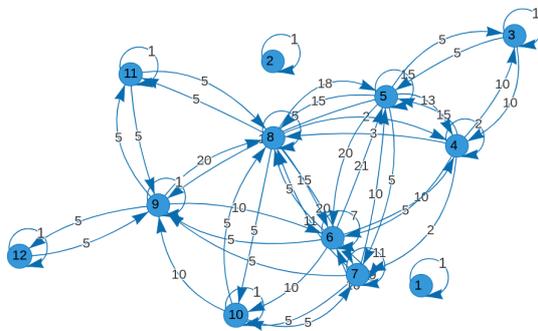
Figura 27 – Exemplo de grafos produzidos a partir do mesmo sinal no domínio do tempo (amostras à esquerda) e no domínio da frequência (amostras à direita). Nota-se uma simplificação da representação comparando as duas perspectivas do mesmo sinal.



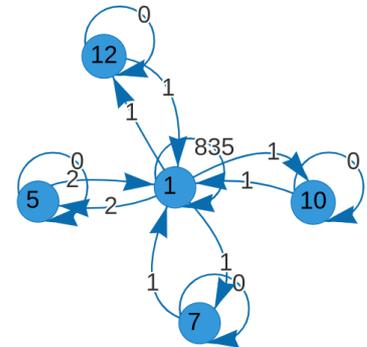
(a)  $L=6$  no tempo.



(b)  $L=6$  na frequência.



(c)  $L=12$  no tempo.



(d)  $L=12$  na frequência.

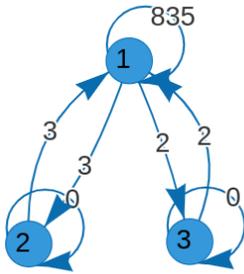
Fonte: Elaborada pelo autor.

as diferentes frequências do sinal que foram impostas pelo ruído adicionado. Como pode ser observado na Figura 26, o primeiro nível representado pelo vértice 1 indica o intervalo que corresponde as magnitudes do ruído, sendo facilmente identificado pelo vértice de maior grau.

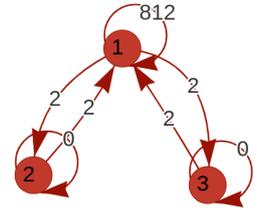
Por outro lado, a estrutura do grafo também representa outra característica topológica do sinal no domínio da frequência. Neste caso, destacamos o ciclo<sup>1</sup> com arestas de maior peso, cujos vértices indicam o nível de magnitude onde há maior concentração de marcadores. No caso da Figuras 28a, que representa o sinal da classe 0, há um ciclo entre os vértices 1 e 2, notoriamente, o nível 2 para  $L = 3$  apresenta maior concentração de marcadores  $\alpha$ , isto é, pontos de máximos locais do sinal, como apresentado na Figura 26. Este mesmo comportamento acontece para  $L=6$ , observe o ciclo entre os vértices 1 e 3, na Figura 28c. Vale destacar que esta análise pode sofrer influências com o nível de granularidade, isto é, o valor de  $L$  utilizado.

<sup>1</sup> Um ciclo em teoria de grafos é um caminho em que o primeiro e o último vértice coincidem.

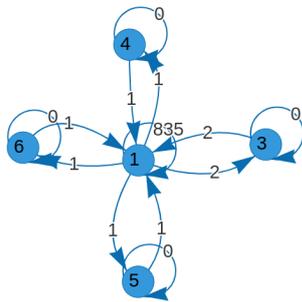
Figura 28 – Exemplo de digrafos criados a partir de duas amostras, uma do sinal da classe 0 (em azul) e uma do sinal da classe 1 (em vermelho). Ambas sob diferentes construções com  $L$  assumindo os valores 3, 6 e 12.



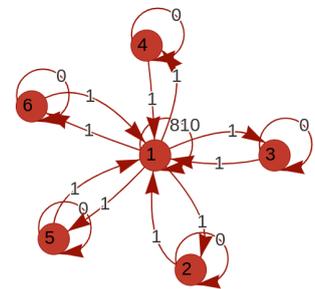
(a)  $L=3$  classe 0



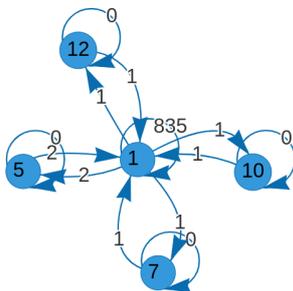
(b)  $L=3$  classe 1



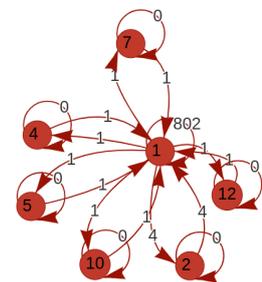
(c)  $L=6$  Classe 0



(d)  $L=6$  Classe 1



(e)  $L=12$  Classe 0



(f)  $L=12$  Classe 1

Fonte: Elaborada pelo autor.

### 4.3 Análise da discriminação de amostras entre classes.

Por outro lado, após o processo de extração de características, é salutar considerarmos o quanto as amostras nas populações do dataset são distintas em suas respectivas classes. Esta preocupação se justifica pelo fato de que quanto maior a separabilidade das classes, tende-se a ter um melhor resultado quando as amostras representativas, resultante dos descritores de características, são submetidas a tarefa de classificação em conjunto com modelos de classificação.

Neste sentido, como descrito na Seção 2.6.1, o teste ANOVA é uma técnica estatística útil para comparar médias entre grupos independentes, apresentando-se como uma valiosa metodologia quando se deseja determinar se há diferenças significativas nas médias entre

populações de amostras, podendo ser aplicada para avaliar a capacidade de um algoritmo de extração de atributos em distinguir grupos de amostras.

Assim, depois de aplicar o descritor de características no sinal, cada digrafo foi representado por um tipo diferente de matriz. Como discutido na Seção 2.1, o grafo pode ser representado sob diferentes perspectivas a partir de diferentes matrizes. Tais matrizes tomam como base os vértices e seus relacionamentos representados nas arestas. Sendo assim, avaliamos individualmente cada matriz Laplaciana Combinatória (CL), Adjacência (A), Laplaciana (L) e Laplaciana Combinatorial Direcionada (DL). Visamos, desta forma, avaliar o comportamento das amostras quanto ao grau de separabilidade na distribuição das amostras.

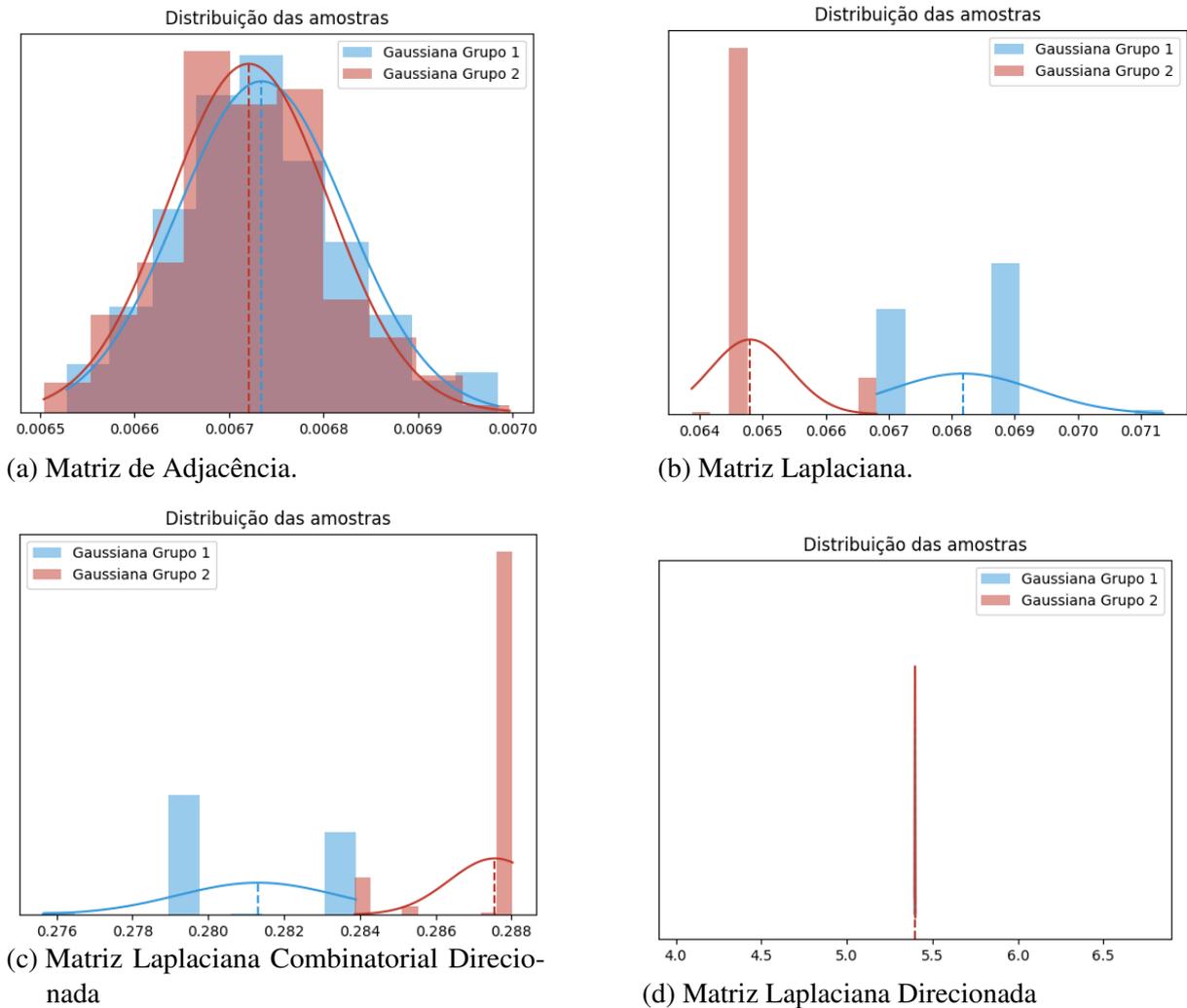
A Figura 29 apresenta diferentes projeções de histogramas das amostras representadas por cada tipo de matriz. A Figura 29-a, apresenta a distribuição das amostras, em azul, amostras da classe 0, em vermelho, amostras da classe 1. É possível notar que os dados apresentam uma característica de distribuição gaussiana, o que indica a possibilidade de uso do teste ANOVA. Neste mesmo caso, nota-se na linha vertical tracejada a média dos grupos. Há uma sutil diferença na média dos dois grupos, entretanto, não sendo assim uma diferença suficientemente satisfatória que possa indicar que apenas a matriz de adjacência seja capaz de fornecer uma boa discriminação entre classes.

Por outro lado, considerando a matriz L, os resultados na Figura 29-b indicam uma melhor separação entre os grupos. É notório, neste caso, que as amostras da classe 1 se concentram na extremidade esquerda do histograma, já as amostras da classe 0, distribuem-se mais à direita. Entre elas é possível perceber a marcação das médias das classes, há uma maior separabilidade entre as curvas gaussianas. Característica interessante para o processo de classificação. O mesmo resultado também foi observado para as amostras representadas pela matriz DL, apresentados na Figura 29-c. Por fim, os casos menos satisfatório quanto ao grau de separabilidade foi apresentado pela matriz CL.

#### **4.3.1 Teste estatístico ANOVA**

Sob outra perspectiva, quanto ao teste estatístico ANOVA, tal análise é balizada pelo nível de significância ( $\alpha = 0.05$ ), quando o valor- $p$  é menor que  $\alpha$ , rejeitamos hipótese nula. Neste sentido, a Figura 30 apresenta quatro diferentes situações, uma para cada tipo de marcador espectral e seu resultado sobre o sinal. Em cada uma das situações é apresentado o valor- $p$ , calculado com base na amostra resultante da combinação entre o número de vértices do digrafo

Figura 29 – Histograma das amostras quanto a sua distribuição e média das classes, de acordo com a extração de atributos representada pelas matrizes de representação do grafo.



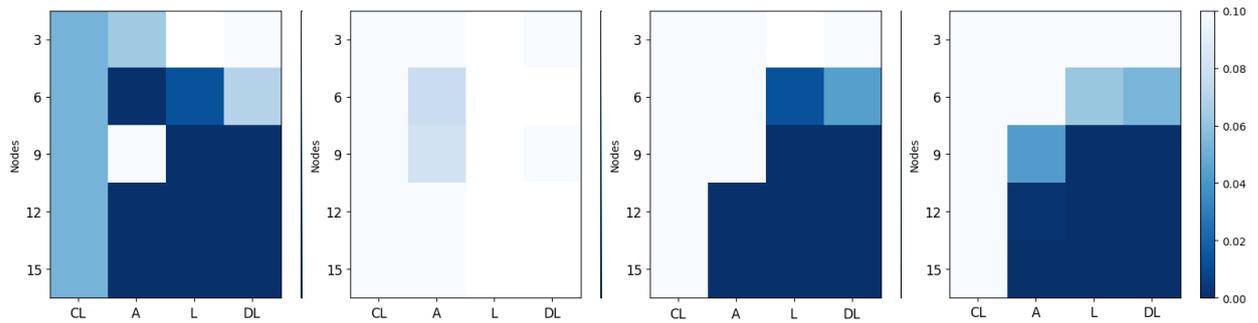
Fonte: Elaborada pelo autor.

e a sua matriz de representação. Observando a escala ao lado, nota-se que quanto mais escura a tonalidade do azul, mais abaixo do nível de significância está a combinação.

A Figura 30-a, apresenta os resultados do marcador espectral  $\alpha$ , que leva em consideração os extremos máximos locais. Nota-se que as matrizes A, L e DL apresentam valor- $p$  menor que o nível de significância para todas as combinações acima de 3 vértices, o que indica a rejeição da hipótese nula, ou seja, há de fato diferenças significativas entre os grupos de amostras e que tais diferenças podem inferir na separação das amostras em populações diferentes.

De outro modo, entretanto, o marcador espectral  $\beta$ , que leva em consideração os extremos mínimos locais, não apresenta valor- $p$  que indique separação suficiente entre classes. Este resultado corrobora com a construção do marcador, uma vez que ele prioriza coeficientes

Figura 30 – Resultado do valor- $p$  no teste ANOVA para discriminação entre as duas classes de sinais. Em cada matriz, as linhas representam o número de vértices e as colunas, mas matrizes de representação do grafo.



(a) Marcador espectral  $\alpha$  (b) Marcador espectral  $\beta$  (c) Marcador espectral  $\gamma$  (d) Marcador espectral  $\delta$

Fonte: Elaborada pelo autor.

espectrais que representam mínimos e, no caso destes sinais, há poucos coeficientes de mínimos locais efetivos, aqueles que existem representam ruídos e não contribuem para a separação das populações de amostras, como pode ser observado na Figura 25.

Já o resultado do marcador espectral  $\gamma$ , que une máximos e mínimos e também considera os pontos de inflexão, apresentou resultados para o valor- $p$  que indicam a rejeição da hipótese nula, ou seja, as amostras representadas por estes marcadores apresentam diferenças que podem inferir na separação das amostras em classes distintas.

Diante dos resultados, considerando o perfil de distribuição apresentado no histograma e teste estatístico ANOVA, é possível inferir que há potencial de separação das amostras de sinais entre classes, quando estas são representadas pelas características do descritor sob a perspectiva das matrizes A, L e DL. O que também sugere que a combinação de tais matrizes pode contribuir para robustez do método quanto a tarefa de classificação.

Neste sentido, a metodologia sugere a combinação das três matrizes A, L e DL em um único tensor, chamado aqui de tensor espectral, como representação das amostras originais. Esta representação se torna atrativa especialmente quanto a combinação com algoritmos recentes para classificação de amostras, que são baseados em tensores e combinam diferentes perspectivas em cada dimensão do tensor, seja por meio da representação de cores, como acontece com a classificação de imagens, seja por meio da representação de vértices e arestas, como na metodologia proposta.

Por fim, a Tabela 3 apresenta três diferentes medidas já consagradas na literatura (HAN *et al.*, 2022; DUDA *et al.*, 2006), a saber: distância de Manhattan e Euclidiana ao quadrado, além do índice de Fisher, já utilizadas quanto a seleção de atributos. Tais medidas

avaliam o potencial de separação entre classe do tensor espectral como um todo, isto é, com as três matrizes A, L e DL combinadas, apresentando medidas de distância tal como Manhattan, Euclidiana e índice de Fisher. Esse último, por exemplo, aglutina medidas intra e inter classes, é uma medida que avalia a separação entre classes com base nas variâncias entre as médias das características das classes e nas variâncias dentro das próprias classes.

Neste sentido, a interpretação do índice de Fisher próximo a 0 indica que as classes estão sobrepostas ou muito próximas uma da outra em termos das características consideradas, já o índice de Fisher maior que 0 indica que quanto maior o valor, melhor é a separação entre as classes em comparação com a variabilidade dentro delas. Portanto, um índice de Fisher de 6,3665 e 7,0720, como aqueles apresentados pelos marcadores  $\alpha$  e  $\gamma$ , respectivamente, sugere que as médias das características entre as duas classes são significativamente distintas em comparação com a variabilidade dentro de cada classe. Isso é positivo se a separação entre as classes for um objetivo importante no contexto do problema.

Vale destacar que os resultados apresentados na Tabela 3, referem-se a medidas considerando, neste caso, o tensor espectral. Os resultados corroboram com os testes ANOVA, apresentados anteriormente, observe que os resultados do marcador  $\beta$  indicaram piores índices de separação, tal como aconteceu nos resultados apresentados na Figura 30, por outro lado, os melhores resultados de indicador de distância de Manhattan, Euclidiana e índice Fisher foram apresentados pelos marcadores  $\alpha$ ,  $\beta$  e  $\delta$ , respectivamente. O que reforça a indicação de que os marcadores propostos nesta tese apresentam potencial de separabilidade entre classes.

#### **4.3.2 Tempo de extração**

Quanto ao tempo necessário para a extração, este experimento executou extrações com diferentes configurações do descritor proposto, variando a construção do digrafo com 3, 6, 9, 12 e 15 vértices, em combinação com os diferentes marcadores espectrais propostos. A tabela 4 apresenta os resultados de tempo médio para 400 diferentes sinais, com duração de 5 segundos, compostos por diferentes frequências e com ruído induzido.

É possível notar que os marcadores  $\alpha$  e  $\delta$  apresentam maior tempo de processamento considerando a configuração com 15 vértices, com tempo médio de 0.008 segundos para cada amostra, o desvio padrão aparece com valor zero já que considerou-se uma precisão até a terceira casa decimal. O maior tempo apresentado para o extrator  $\delta$ , por exemplo, já era esperando quando comparado com o marcador espectral  $\beta$ , uma vez que os sinais avaliados apresentam

Tabela 3 – Relação entre marcadores, número de vértices e medidas de distância entre as classes considerando o tensor espectral proposto.

Marcador	Nodes	Dist. Manhattan	Dist. Euclidiana ao Quadrado	Índice de Fisher
Marcador espectral $\alpha$	3	0,3951	0,1561	<b>6,3665</b>
	6	<b>2,4087</b>	<b>5,8020</b>	2,8500
	9	0,8239	0,6787	6,2882
	12	0,4607	0,2122	0,8338
	15	0,1567	0,0245	0,0321
Marcador espectral $\beta$	3	<b>0,0116</b>	<b>0,0001</b>	<b>0,6460</b>
	6	0,0116	0,0001	0,6460
	9	0,0116	0,0001	0,6460
	12	0,0116	0,0001	0,6460
	15	0,0116	0,0001	0,6460
Marcador espectral $\gamma$	3	0,3950	0,1560	6,3418
	6	<b>2,6712</b>	<b>7,1352</b>	2,9055
	9	0,8646	0,7475	<b>7,0721</b>
	12	0,5945	0,3535	1,2792
	15	0,1666	0,0278	0,0322
Marcador espectral $\delta$	3	0,0231	0,0005	0,6810
	6	0,0231	0,0005	0,6810
	9	0,0231	0,0005	0,6810
	12	0,0231	0,0005	<b>0,6810</b>
	15	<b>0,1945</b>	<b>0,0378</b>	0,0042

Fonte: Elaborada pelo autor.

Tabela 4 – Tempo médio de extração para 400 sinais com duração de 5 segundos, composto por diferentes frequências e com ruído induzido.

Tipo do Marcador espectral	Vértices	Tempo médio (s)
Marcador espectral $\alpha$	3	<b>0.003</b>
	6	0.004
	9	0.005
	12	0.005
	15	0.008
Marcador espectral $\beta$	3	0.003
	6	0.003
	9	0.004
	12	0.004
	15	0.005
Marcador espectral $\gamma$	3	0.003
	6	0.004
	9	0.005
	12	0.005
	15	0.006
Marcador espectral $\delta$	3	0.004
	6	0.004
	9	0.004
	12	0.005
	15	<b>0.008</b>

Fonte: Elaborada pelo autor.

uma maior quantidade de frequências com picos de magnitudes em relação a quantidade de coeficientes de magnitudes em pontos de mínimos do sinal, no domínio da frequência.

#### 4.4 Conclusão do Capítulo

Neste capítulo, exploramos os resultados obtidos pela metodologia proposta em dois conjuntos distintos de sinais, cada conjunto formado por 200 amostras. O primeiro conjunto

consiste em sinais senoidais, formados por cinco frequências, enquanto o segundo foi produzido por dez frequências. Em ambos os casos, introduzimos ruídos gaussianos para simular condições realísticas de aquisição. Após a extração de atributos, apresentamos as distribuições das amostras por meio de histogramas e suas curvas gaussianas, além do cálculo da distância entre classes, com base nas medidas de Manhattan, Euclidiana e índice Fisher. Foi possível observar uma separação entre classes bastante promissora, especialmente para as matrizes A, L e DL.

Também foi possível avaliar a separabilidade média entre os grupos por meio do teste estatístico ANOVA. Considerando grafos com 3, 6, 9, 12 e 15 vértices e suas respectivas matrizes. Comparamos os valores- $p$  obtidos em cada combinação, observou-se que as matrizes apresentam um valor- $p$  que rejeita a hipótese nula, o que indica potencial de representação das amostras, rejeitando a hipótese nula em pelo menos umas das matrizes, permitindo inferências significativas na separação das amostras em classes distintas, indicando seu potencial para classificação.

## 5 RESULTADOS EM CONTEXTO MONOCANAL: CLASSIFICAÇÃO DA DOENÇA DE HUNTINGTON VIA SINAIS DE VOZ.

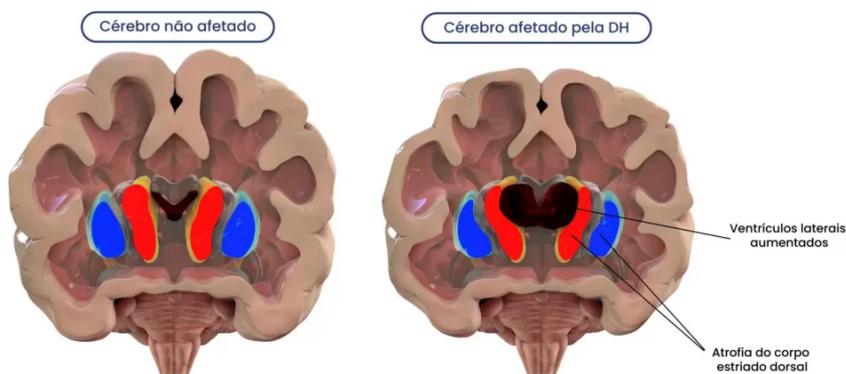
Neste capítulo apresentamos uma aplicação do descritor de sinais em amostras de pacientes saudáveis e diagnosticados com a doença de Huntington. Nesta aplicação utilizou-se sinais de voz, representando uma aplicação monocanal do descritor. Os tensores espectrais produzidos representam os pacientes e são utilizados na fase de classificação.

### 5.1 A doença de Huntington

A doença de Huntington (HD, do inglês *Huntington's disease*) caracteriza-se como uma condição hereditária em que as células nervosas do cérebro se rompem ao longo do tempo. Esta degradação das células causam distúrbios cerebrais progressivos, contemplando também problemas de ordem emocionais, incidências na perda da capacidade de raciocínio, além de quadros de amnésia, confusão mental, delírio, falta de concentração, lentidão durante atividades ou dificuldade em pensar e compreender, além de outros distúrbios cerebrais.

Estudos recentes investigam a relação entre mutação genética e as causas desta doença com base no estudo do gene HTT, que fornece instruções para a síntese da proteína huntingtina. A proteína alterada constrói aglomerados nas células cerebrais ao longo do tempo, causando dano e morte neuronal e uma das áreas do cérebro mais comumente afetadas pela HD é o corpo estriado dorsal, região responsável pelo controle parcial do movimento voluntário e de funções cognitivas como a memória de trabalho (CERVATO, 2023). A Figura 31 apresenta uma comparação do cérebro saudável em relação as regiões de um cérebro portador da doença.

Figura 31 – Principais degradações no corpo estriado dorsal do cérebro que afetam o controle parcial do movimento voluntário. Adaptado de Cervato (2023).



Fonte: Elaborada pelo autor.

Como a doença apresenta uma característica hereditária, uma criança tem meio por

cento de probabilidade de herdar uma cópia do cromossomo que transmite a doença. Atualmente não existe cura ou medicamento que possa parar ou reverter a doença de um paciente, mas existem tratamentos capazes de controlar os sintomas da doença e o diagnóstico precoce contribui para a efetividade do tratamento. O paciente também pode usar medicamentos e terapia direcionados ao tratamento da depressão para aliviar os sintomas (IRAM *et al.*, 2015).

Os indivíduos afetados com a doença apresentam gradativamente dificuldades para falar e engolir, devido aos danos nos movimentos involuntários, além da incidência de alterações de voz e na fala, sintomas que são fortemente correlacionados com esta doença. Deste modo, visando uma proposta financeiramente mais acessível e com tempo de resposta reduzido, trabalhos recentes se dedicam ao estudo dos efeitos na fala e na voz dos pacientes, atuando sob o ponto de vista da análise acústica da voz para diagnosticar a doença (GARCÍA *et al.*, 2011; GALLEZOT *et al.*, 2022).

A metodologia apresentada neste capítulo, portanto, propõe uma abordagem unindo técnicas clássicas e recentes para extração de características presentes nos sinais produzidos pelos pacientes, com o objetivo de classificar amostras com base em sinais de voz. Deste modo, os sinais são produzidos a partir do registro da voz durante a leitura de dois poemas lituanos. A partir destas amostras, tecnologias clássicas como a transformada de Fourier combinada com técnicas da teoria dos grafos tornam possível a identificação das componentes de magnitude do sinal, conhecidas também como componentes espectrais. A leitura dos sinais de voz e sua representação sob estruturas baseadas em Teoria dos Grafos possibilita uma análise de componentes espectrais do domínio da frequência sob uma perspectiva topológica, representando o sinal por meio de vértices e arestas ponderadas.

## **5.2 Trabalhos relacionados a detecção da doença**

No capítulo inicial desta tese, especificamente na Seção 1.3.1.1, apresentamos diferentes abordagens para extração de atributos em análise de sinais monocanais que contribuem para a pesquisa no âmbito deste capítulo, além de apresentar os trabalhos relacionados que são utilizados com fins comparativos nos resultados desta aplicação.

### 5.3 Amostras de pacientes

O conjunto de dados utilizado neste estudo foi compilado através da captação de gravações áudio de indivíduos saudáveis e diagnosticados com a doença. Cada participante foi orientado a recitar dois poemas da autoria de Maironis, um poeta lituano. Estas duas obras poéticas foram "*Lietuva brangi*" e "*Trakų pilis*". o conjunto de dados abrange um total de 120 gravações de fala dos participantes saudáveis e 66 gravações de fala dos participantes afetados pela doença de Huntington. Uma vez captadas as gravações, estas foram convertidas para os formatos de áudio WAVEform e MPEG-4 para posterior análise e processamento. Os paciente com diagnóstico da HD foram categorizado de acordo com o sistema de avaliação Shoulson-Fahn, sendo classificados nos estágios iniciais da doença, I e II.

### 5.4 Metodologia

A construção do digrafo para a representação das amostras segue exatamente a metodologia descrita no Capítulo 3. Na seção a seguir, descrevemos a etapa de aquisição do sinal e extração dos atributos. Por fim, ao término da seção, apresentamos as configurações de ambiente e metodologia de avaliação do desempenho dos modelos.

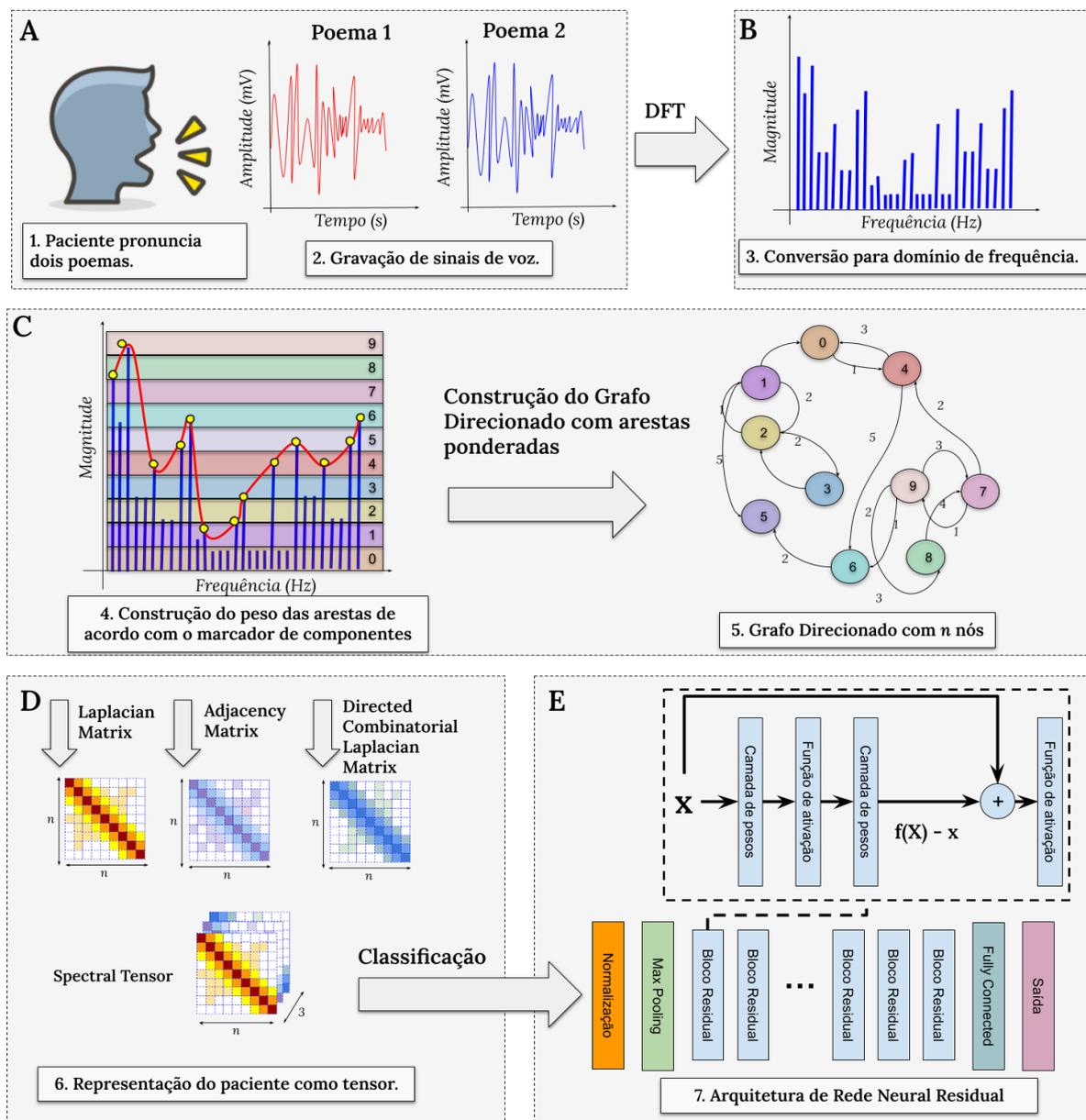
#### 5.4.1 Construção do Digrafo

A Figura 32 mostra um processo passo-a-passo de extração de características do sinal de voz produzido pelos doentes. Inicialmente, cada doente pronuncia dois poemas lituanos e é gravado um sinal áudio para cada poema. Esta fase representa a aquisição do sinal. Após a leitura dos sinais, seguem as fases de conversão para o domínio da frequência, Figura 32-B, seguida da identificação dos marcadores espectrais, Figura 32-C, e por fim a construção do digrafo e a combinação das matrizes no tensor espectral, Figura 32-D.

#### 5.4.2 Avaliação de Desempenho

Os diferentes marcadores de componentes espectrais foram combinados com diferentes arquiteturas de redes neurais, e assim avaliados quanto ao seu desempenho sob o ponto de vista das métricas de Acurácia, Sensibilidade, a Especificidade e F1-score. Para construir o grafo direcionado foram avaliados diferentes níveis  $L$  considerando a estratégia de busca em grade,

Figura 32 – Metodologia para a extração de características de sinais de voz, passos A-D. Nesta proposta, os pacientes são representados pelo tensor espectral do sinal.



Fonte: Elaborada pelo autor.

onde o número de nós avaliados variou entre 12 e 76, com um passo de 4. Cada valor de  $L$  foi avaliado de acordo com o valor de acurácia apresentado pela fase de classificação dos tensores.

A fase de classificação dos tensores espectrais gerados pela metodologia de extração considerou a estratégia de validação cruzada com amostragem estratificada, com  $K=5$ . Neste sentido, o modelo de predição usa  $K - 1$  dobras para treinamento e reserva o restante para teste. Este procedimento é repetido duas vezes seguidas para seleção aleatória das amostras. Todos os

modelos de classificação apresentados no Capítulo 2 utilizaram o conjunto de treino e teste de tensores espectrais.

## 5.5 Resultados e discussão

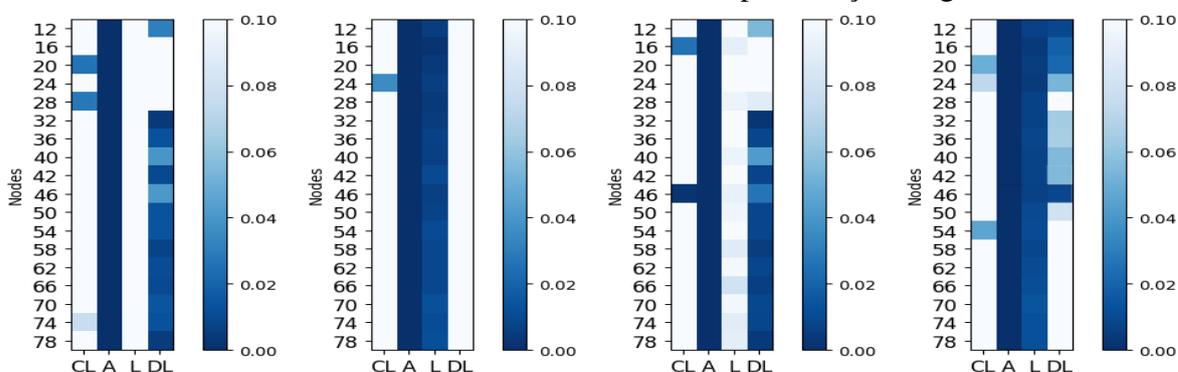
Os valores alcançados pela metodologia proposta neste trabalho são avaliados sob duas perspectivas: O desempenho do método proposto em relação aos marcadores de componentes espectrais e a relevância dos resultados da classificação sobre extratores de atributos clássicos e extratores especialistas em áudio. O ambiente computacional utilizado possui as seguintes configurações: Intel Core i7 2.2GHz, 16GB de RAM, Python 3.7, Torch 1.10, Scikit-learn 1.0.2.

### 5.5.1 Teste estatístico ANOVA

Sob a perspectiva do teste estatístico ANOVA, a hipótese nula ( $H_0$ ) considerada é que não haja separabilidade significativa entre os grupos de amostras, já a hipótese alternativa ( $H_1$ ) indica que há de fato diferenças significativas e que tais diferenças podem inferir na separação das amostras em populações diferentes. Neste sentido, a Figura 33 apresenta quatro diferentes resultados, uma para cada tipo de marcador espectral e seu resultado sobre o sinal. Observando a escala ao lado, representado o valor- $p$ , nota-se que quanto mais escura a tonalidade do azul, mais abaixo do nível de significância ( $\alpha = 0.05$ ) está a combinação.

A Figura 33-d apresenta os resultados do marcador espectral  $\delta$ , que leva em consideração os pontos de inflexão do sinal. Nota-se que as matrizes A, L e DL apresentam valor- $p$

Figura 33 – Resultado do valor- $p$  no teste ANOVA para discriminação entre as duas classes de sinais. Em cada matriz, as linhas representam o número de vértices e as colunas, mas matrizes de representação do grafo.



(a) Marcador espectral  $\alpha$  (b) Marcador espectral  $\beta$  (c) Marcador espectral  $\gamma$  (d) Marcador espectral  $\delta$

Fonte: Elaborada pelo autor.

menor que o nível de significância para a maioria das combinações. A tonalidade apresentada na figura indica a rejeição da hipótese nula, ou seja, há de fato diferenças significativas entre os grupos de amostras e que tais diferenças podem inferir na separação das amostras em populações diferentes.

Vale também ressaltar que a matriz  $A$  e matriz  $L$  apresentaram maior potencial de separação entre as classes, considerando o valor- $p$  significativamente abaixo do nível de confiança, para os marcadores espectrais  $\beta$  e  $\delta$ . Estes valores indicam que no casos destas amostras de sinais, há maior relevância para separação das amostras em níveis de magnitude que representam mínimos locais.

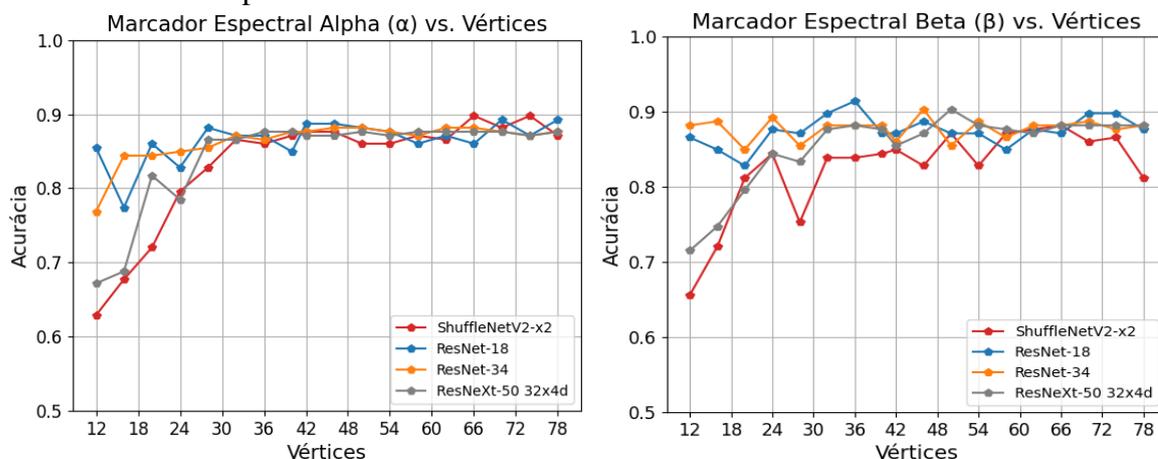
### 5.5.2 Comparação de marcadores de componentes espectrais quanto a classificação

A Figura 34a mostra as taxas de Precisão para o marcador de coeficiente mínimo local  $\alpha$ . Neste sentido, entre os 18 modelos avaliados, são apresentados os quatro melhores modelos: ShuffleNetv2-x2 ( $89,78 \pm 4,45$ ), ResNet-18 ( $89,25 \pm 5,51$ ), ResNet-34 ( $88,17 \pm 6,73$ ) e ResNeXt-50 32x4d ( $87,63 \pm 7,49$ ). É possível observar na Figura 34a que um menor número de nós para análise de sinais apresentou menores taxas de precisão.

Esse comportamento pode indicar que um conjunto de menor número de nós tem dificuldades em representar melhor os marcadores entre os diferentes nós do grafo direcionado e suas respectivas arestas ponderadas. O menor número de nós, portanto, prejudicou a representação das amostras. Esse mesmo resultado pode ser observado nos marcadores  $\beta$ ,  $\gamma$  e  $\delta$ , Figura 34a. O marcador  $\alpha$  demonstrou as menores acurácias para todas as combinações de classificadores e o número de nós no grafo avaliado. Nenhuma das combinações apresentou uma precisão superior a 90%.

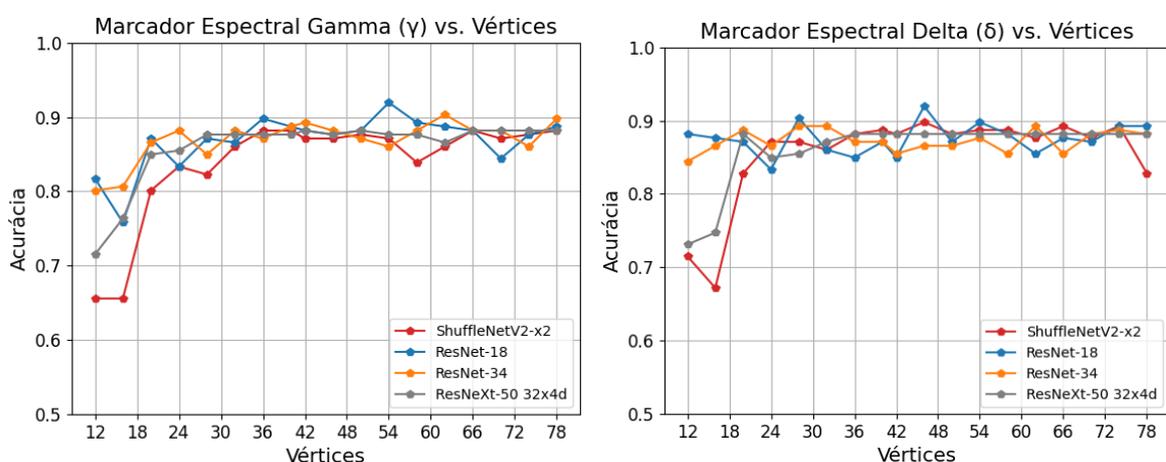
Por outro lado, o marcador  $\beta$ , que representa os coeficientes de máxima magnitude local, apresentou uma melhora significativa no desempenho para a configuração  $L = 36$ , apresentando o melhor resultado para este marcador com uma Acurácia de  $91,40 \pm 02,17$ . Acredita-se que o melhor resultado, em relação ao marcador  $\alpha$ , deve-se ao número mais significativo de marcadores, uma vez que o sinal de voz apresenta uma maior quantidade de componentes de frequência de baixa magnitude no sinal. Essa maior quantidade afeta o peso nas arestas do grafo direcionado, afetando diretamente a representação do tensor espectral dos pacientes. Nesse sentido, é possível concluir que o número mais significativo de marcadores contribuiu para a melhor representação das amostras. A Figura 34b apresenta estes resultados para o marcador  $\beta$ .

Figura 34 – Taxas de acerto de acordo com o número de nós para os marcadores espectrais



(a) Marcador espectral  $\alpha$

(b) Marcador espectral  $\beta$



(c) Marcador espectral  $\gamma$

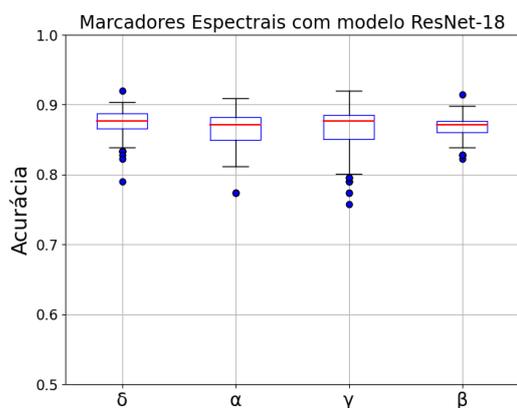
(d) Marcador espectral  $\delta$

Fonte: Elaborada pelo autor.

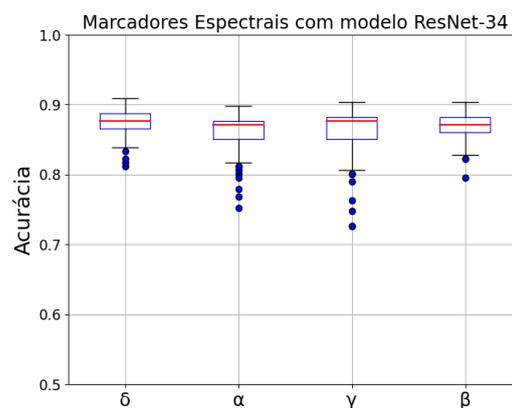
O marcador  $\gamma$ , que representa a união dos marcadores  $\alpha$  e  $\beta$ , por outro lado, apresentou resultados acima de 90% de precisão, considerando  $L = 54$  com o modelo ResNet-18 atingindo  $91,94 \pm 01,40$ , com resultados de Precisão muito promissores atingindo  $92,00 \pm 01,31$ , indicando uma taxa de sucesso na identificação de pacientes com doença de Huntington.

Destaca-se o marcador definido pelos pontos de inflexão do sinal, o marcador  $\delta$ . Este marcador representa, para além dos pontos de máximo e mínimo, os pontos onde ocorrem alterações no crescimento e diminuição do sinal. O marcador  $\delta$  apresentou estabilidade mais significativa em relação à variação do número de nós do gráfico. Destacando resultados já próximos de 90% de acurácia para o modelo com  $L = 12$  e a rede ResNet-18, com  $88,17 \pm 14,51$ , uma diferença de mais de 7% em relação à mesma combinação para o marcador  $\gamma$ , que atingiu  $81,72 \pm 20,42$ .

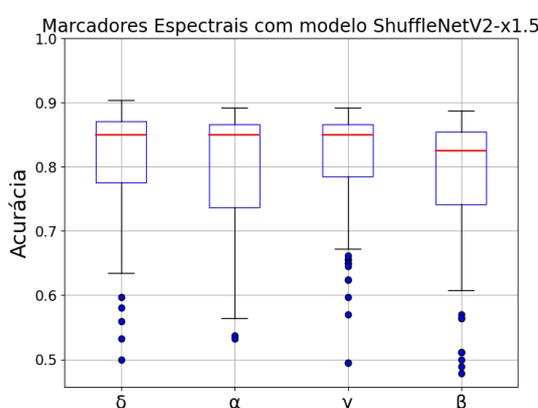
Figura 35 – Top 4 melhores modelos de classificação com base nas taxas de acerto



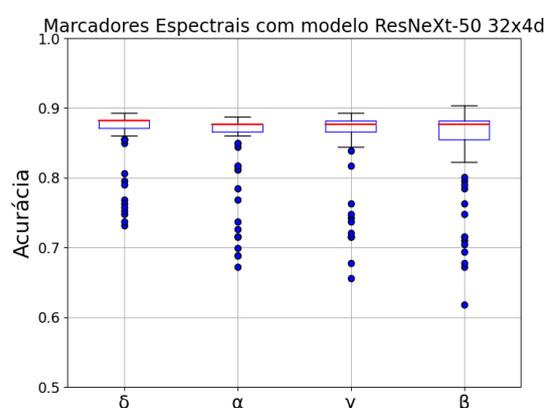
(a) ResNet-18



(b) ResNet-34



(c) ShuffleNetV2-x1.5.



(d) ResNeXt-50 32x4d

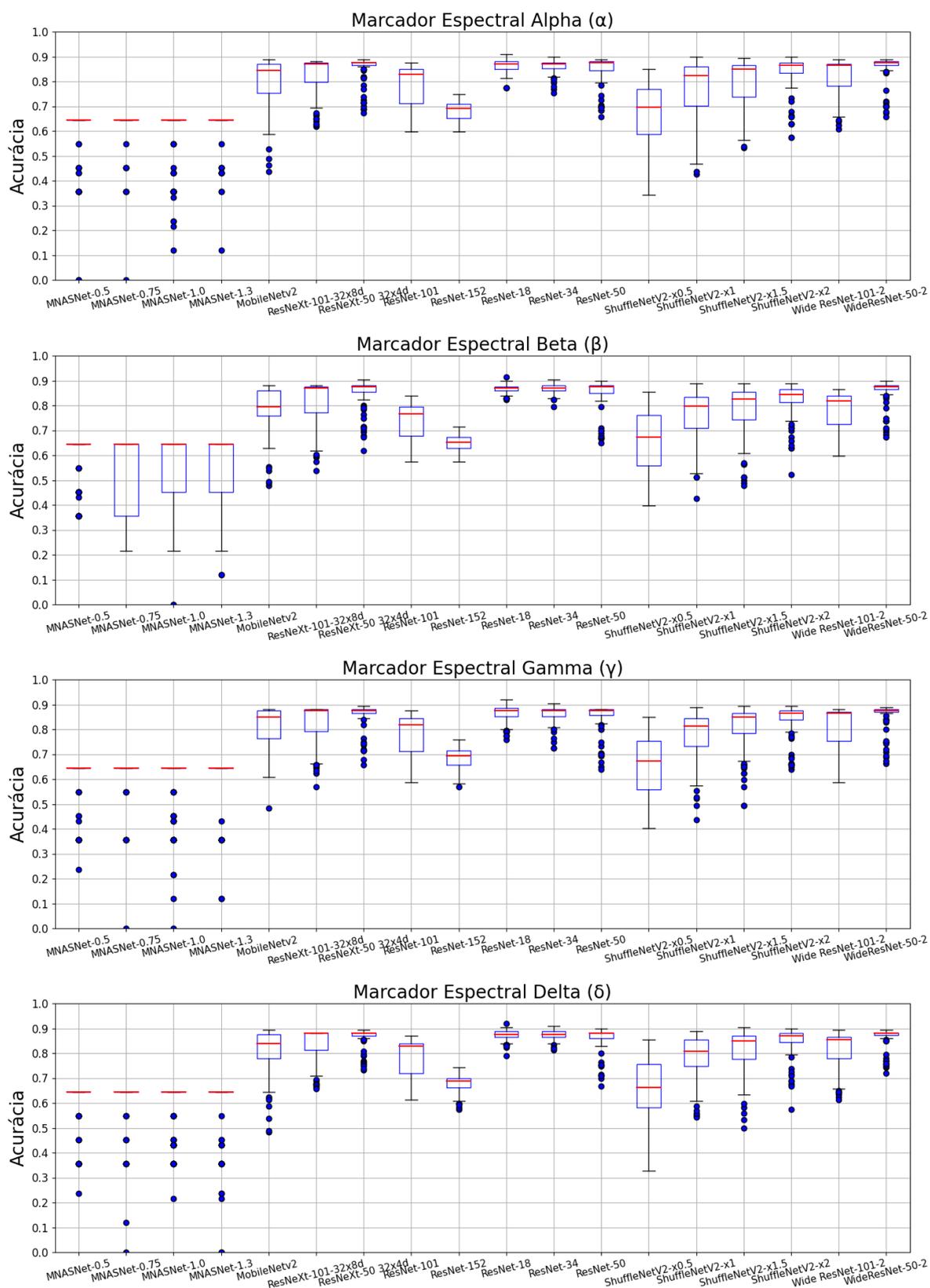
Fonte: Elaborada pelo autor.

Por fim, cada marcador proposto neste estudo também foi avaliado quanto à estabilidade dos resultados apresentados sobre as taxas de acerto combinadas com o modelo de classificação. A Figura 35d apresenta os resultados para as quatro melhores arquiteturas. O modelo ResNet-18 obteve melhores medianas, com baixa oscilação entre os *outliers*. O marcador com as flutuações mais significativas nos resultados foi o  $\gamma$ , apresentando muitos *outliers*, como mostra a Figura 35d para o marcador  $\delta$ .

O marcador  $\delta$  combinado com a arquitetura ResNet-18 apresentou resultados estáveis relativamente a variações no número de nós no gráfico, Figura 36. Além disso, essa combinação sugere índices promissores de Acurácia ( $91,94 \pm 01,40$ ), Recall ( $91,94 \pm 01,40$ ) e Precisão ( $92,00 \pm 01,31$ ) para classificação entre pacientes saudáveis e pacientes com doença de Huntington.

A Figura 36 apresenta os resultados gerais para cada um dos marcadores propostos.

Figura 36 – Desempenho dos marcadores  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$  de acordo com o índice médio de Acurácia no conjunto de teste.



Fonte: Elaborada pelo autor.

Alguns modelos de classificação derivados da família de redes neurais MnasNet e ShuffleNet apresentaram as variações mais significativas nas taxas de acerto, com precisões medianas entre 60% e 70% e abaixo de 90%, respetivamente. Essas variações não são benéficas para a tarefa de classificação de padrões, pois podem indicar uma maior dependência do modelo sobre a combinação de seus parâmetros e dados de entrada, indicando baixa capacidade de generalização. Por outro lado, os modelos derivados da arquitetura ResNet apresentaram uma menor distribuição de resultados, indicando uma mediana muito próxima dos 90%. Alguns casos ultrapassam 91% de precisão, como nos modelos ResNet-18 e ResNet-34.

### 5.5.3 Comparação com extratores de atributos clássicos e wavelets

Visando ampliar a avaliação, também considerou-se os resultados dos extratores clássicos com foco na extração de características de sinais unidimensionais. Os extratores clássicos incluem Fourier, Ceptral, as diferentes famílias Wavelets e o Statistical Extractor. Esta avaliação considera diferentes propostas de extratores com modelos de classificação apresentados na Tabela 5. Nesta mesma tabela são apresentados os parâmetros selecionados via metodologia RandomizedSearchCV<sup>1</sup>, com 20 iterações e com K=5.

Tabela 5 – Lista dos diferentes modelos de classificação e seus respectivos espaços de parâmetros avaliados para otimização. Os parâmetros referem-se à terminologia do framework.

Classificador	Parâmetros
KNN	'n_neighbors': [3, 5, 7], 'weights': ['uniform', 'distance'], 'p': [1, 2]
RF	'n_estimators': [100, 200, 300], 'max_depth': [None, 5, 10], 'min_samples_split': [2, 5, 10]
ETrees	'n_estimators': [100, 200, 300], 'max_depth': [None, 5, 10], 'min_samples_split': [2, 5, 10]
HGB	'learning_rate': [0.001, 0.01, 0.1, 0.2, 0.3], 'max_depth': sp_randint(3, 10), 'min_samples_leaf': sp_randint(1, 10)
XGB	'learning_rate': [0.001, 0.01, 0.1, 0.2, 0.3], 'max_depth': sp_randint(3, 10), 'min_child_weight': sp_randint(1, 10), 'subsample': [0.6, 0.7, 0.8, 0.9, 1.0], 'colsample_bytree': [0.6, 0.7, 0.8, 0.9, 1.0], 'gamma': [0, 0.1, 0.2, 0.3, 0.4]
MLP	'hidden_layer_sizes': [(50,), (100,), (50, 50), (100, 100), (100, 50, 25)], 'activation': ['relu', 'tanh'], 'alpha': [0.0001, 0.001, 0.01], 'learning_rate': ['constant', 'invscaling', 'adaptive']
SVM	'C': [0.1, 1, 10], 'kernel': ['linear', 'rbf', 'poly', 'sigmoid'], 'gamma': ['scale', 'auto'], 'probability': [True]
Gaussian MNB	'alpha': [0.1, 0.5, 1.0]
Gaussian NB	-
LDA	-
QDA	-
Deep Models	Cross-Entropy Cost Function, AdaGrad, 'learning rate': 0.0001, 30 epochs

Fonte: Elaborada pelo autor.

<sup>1</sup> Busca aleatória para encontrar uma amostra representativa das combinações de hiperparâmetros

De acordo com os resultados apresentados na Tabela 6, os extratores baseados em Wavelet apresentaram dificuldades na classificação de amostras entre pacientes saudáveis e pacientes com Huntington. As taxas de Precisão, Sensibilidade e Recall ficaram próximas de 68%, exceto para o extrator COIF13, que alcançou 72% de acurácia com 76% de precisão. Por outro lado, extratores estatísticos e OHS, embora com número de atributos significativamente menor que aqueles baseados em wavelets, apresentaram resultados mais expressivos, acima de 70% de acurácia, e com 72% para o extrator estatístico. Este resultado indica que um número mais significativo de atributos não corresponde a um bom resultado.

Os resultados mais significativos apresentados pelos classificadores clássicos foram provenientes do extrator de Coeficientes de Magnitude e o extrator MFCC. A melhor combinação apresentada para o extrator MFCC foi com o classificador KNN, atingindo 76% de acurácia e 77% de precisão. Quanto ao extrator MFCC, em sua construção é utilizada uma escala logarítmica que reposiciona as faixas de frequência. A etapa logarítmica pode impor um agrupamento das componentes espectrais, e acredita-se que por este motivo pode haver uma redução na representação das componentes.

Tabela 6 – Comparação entre extratores clássicos na análise de sinal com a abordagem proposta

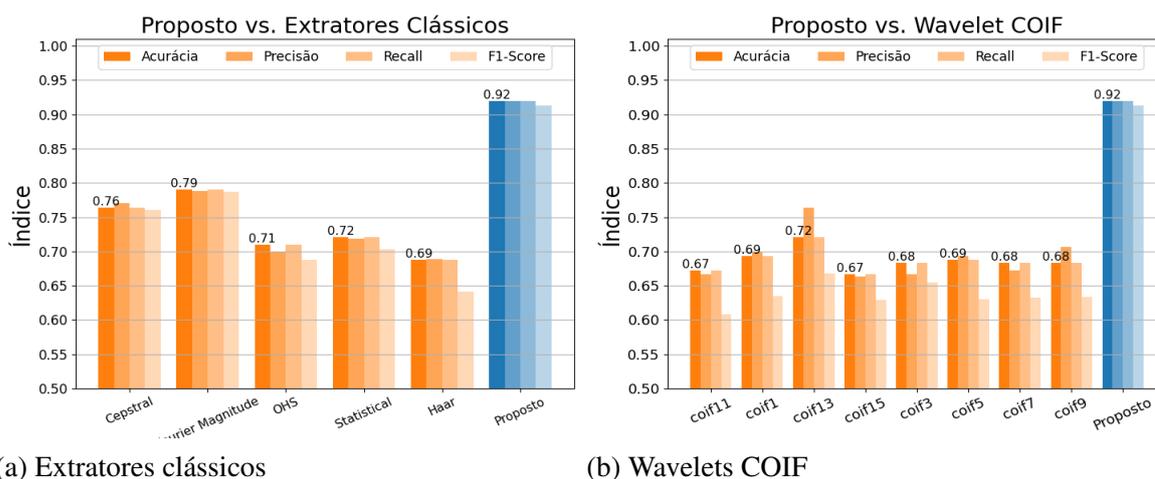
Extrator de áudio	Classificador	Acurácia	Precisão	Recall	F1-Score	Tempo Treino	Tempo Predição
<b>Proposed (δ-46 nodes)</b>	<b>ResNet-18</b>	<b>91,94±01,40</b>	<b>92,00±01,31</b>	<b>91,94±01,40</b>	<b>91,31±02,29</b>	<b>155,8</b>	<b>1,57</b>
MFCCs	KNN	76,34±03,04	77,09±03,16	76,34±03,04	76,05±02,35	42.53	0.31
Coef. de Magnitude	KNN	79,03±03,95	78,81±04,11	79,03±03,95	78,75±03,90	43.36	3.57
OHS	XGB	70,97±02,28	69,95±02,82	70,97±02,28	68,77±03,58	15.37	00.05
Atributos Estatísticos	ETrees	72,04±04,02	71,90±05,81	72,04±04,02	70,32±05,52	03.89	00.02
Wave. COIF11	NB Gaussiano	67,20±03,04	66,63±06,59	67,20±03,04	60,83±03,30	01.48	45.74
Wave. COIF1	NB Gaussiano	69,35±02,28	70,01±02,11	69,35±02,28	63,47±04,43	01.49	35.36
Wave. COIF13	XGB	72,04±00,76	76,37±03,11	72,04±00,76	66,73±01,52	1814.03	48.03
Wave. COIF15	RF	66,67±02,01	66,29±02,65	66,67±02,01	62,91±02,07	36.24	50.06
Wave. COIF3	RF	68,28±00,76	66,62±00,92	68,28±00,76	65,46±01,30	36.53	37.18
Wave. COIF5	NB Gaussiano	68,82±02,01	69,37±04,44	68,82±02,01	62,99±02,15	01.49	39.60
Wave. COIF7	ETrees	68,28±04,02	67,22±06,93	68,28±04,02	63,24±06,09	49.64	41.39
Wave. COIF9	ETrees	68,28±03,80	70,61±08,29	68,28±03,80	63,32±02,60	49.57	43.82
Wave. DB10	NB Gaussiano	67,74±02,63	68,55±05,54	67,74±02,63	61,60±02,48	49.08	02.95
Wave. DB1	ETrees	67,74±03,48	67,45±06,33	67,74±03,48	63,04±02,88	49.08	02.95
Wave. DB15	NB Gaussiano	68,82±02,74	68,53±03,98	68,82±02,74	62,80±04,36	01.50	09.20
Wave. DB20	NB Gaussiano	68,82±04,02	67,78±06,20	68,82±04,02	63,20±05,77	01.50	05.31
Wave. DB25	XGB	68,82±00,76	72,57±05,42	68,82±00,76	62,28±01,29	1894.75	12.94
Wave. DB30	NB Gaussiano	68,28±04,99	66,74±09,75	68,28±04,99	62,84±06,64	01.50	15.23
Wave. DB35	ETrees	68,28±05,32	66,48±08,87	68,28±05,32	61,53±08,04	49.62	17.07
Wave. DB5	NB Gaussiano	68,82±02,74	68,81±05,25	68,82±02,74	63,30±03,78	01.50	05.31
Wave. Haar	ETrees	68,82±00,76	68,87±02,74	68,82±00,76	64,10±01,53	48.67	00.84
Wave. SYM10	NB Gaussiano	67,74±06,03	64,90±09,73	67,74±06,03	61,71±08,40	01.49	28.06
Wave. SYM12	ETrees	68,82±04,62	65,61±09,19	68,82±04,62	63,46±08,02	49.38	29.47
Wave. SYM14	ETrees	68,28±04,99	70,13±10,91	68,28±04,99	62,19±05,01	49.12	31.23
Wave. SYM16	NB Gaussiano	67,20±02,74	66,05±06,16	67,20±02,74	61,04±04,06	01.48	33.26
Wave. SYM2	ETrees	69,35±01,32	69,94±02,08	69,35±01,32	64,43±03,75	49.69	19.45
Wave. SYM4	NB Gaussiano	69,35±01,32	70,27±03,32	69,35±01,32	63,85±01,01	01.50	21.82
Wave. SYM6	NB Gaussiano	68,28±03,80	65,38±08,50	68,28±03,80	62,05±06,86	01.49	23.57
Wave. SYM8	NB Gaussiano	67,74±02,28	67,61±05,23	67,74±02,28	61,19±02,79	01.50	25.70

Fonte: Elaborada pelo autor.

A redução na representação do extrator MFCC pode explicar a diminuição na taxa de acerto dos Coeficientes de Magnitude. Nesse sentido, entre os extratores clássicos, os coeficientes de magnitude apresentam os melhores resultados, apresentando índices relevantes de Precisão ( $79,03 \pm 03,95$ ), Precisão ( $78,81 \pm 04,11$ ) e Recall ( $79,03 \pm 03,95$ ). Em contrapartida, a metodologia proposta também considera as componentes espectrais assim como os extratores MFCC e Coeficientes de Magnitude, porém, sob uma representação de tensores espectrais via representação de um digrafo.

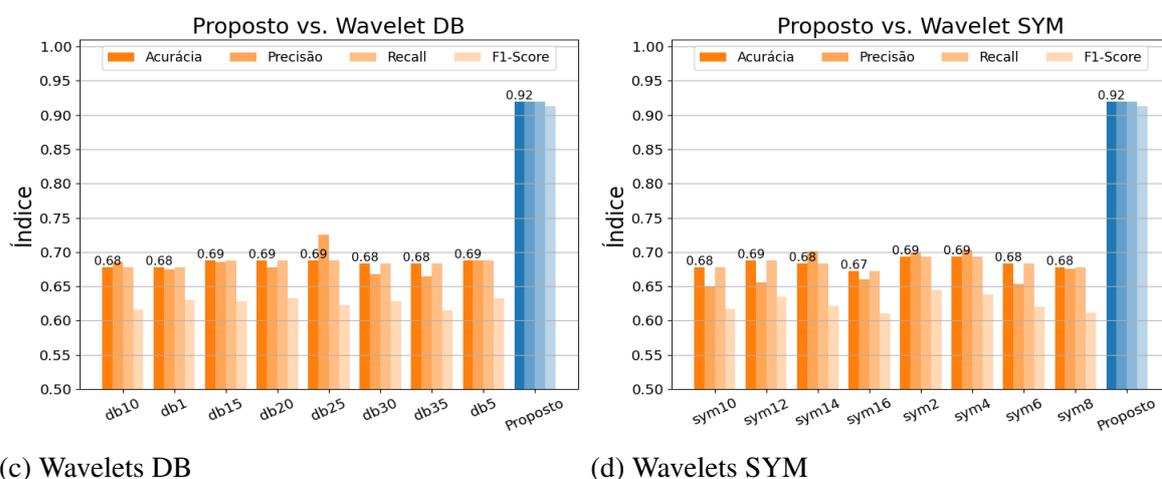
A Figura 37a apresenta diferentes resultados de acurácia, precisão, recall e F1-score para os extratores clássicos. Apenas o Coeficiente de Magnitude apresentou resultados mais expressivos, alcançando 79% de acurácia. A metodologia proposta, no entanto, alcança resultados em torno de 92%.

Figura 37 – Resultados do extratores clássicos e baseados em wavelets.



(a) Extratores clássicos

(b) Wavelets COIF



(c) Wavelets DB

(d) Wavelets SYM

Fonte: Elaborada pelo autor.

Por outro lado, uma comparação com as famílias Wavelet DB (Figura 37c), COIF (Figura 37b) e SYM (Figura 37d) também apresenta uma diferença significativa entre os extratores e o método proposto. Os resultados das famílias SYM e DB não apresentam resultados superiores, com métricas de acurácia média próximas de 68%. A família COIF, por sua vez, apresentou resultados mais promissores com taxas de acurácia acima de 72% combinada com os modelos de classificação XGB. Os resultados destes extratores, embora apresentem um elevado número de atributos após a extração, conforme descrito na Seção 2.3, o que leva a um tempo de treinamento caro para o processo de classificação de doenças, não se mostraram satisfatórios para esta aplicação.

#### 5.5.4 Comparação com extratores de áudio

Por outro lado, a Tabela 7 apresenta diferentes extratores de características especializados em sinais de áudio. A Tabela mostra a combinação dos diferentes modelos com os extratores de características que alcançaram os melhores índices de acurácia entre a classificação de amostras saudáveis e patológicas. Dentre todos os classificadores apresentados na Tabela 5, a Figura 38 mostra todas as melhores combinações da metodologia proposta.

Tabela 7 – Comparação entre extratores especializados em análise de áudio em relação à abordagem proposta

Extrator de Áudio	Classificador	Acurácia	Precisão	Recall	F1-Score	Tempo treino	Tempo Predição
<b>Proposed (<math>\delta</math> - 46 nodes)</b>	<b>ResNet-18</b>	<b>91,94±01,40</b>	<b>92,00±01,31</b>	<b>91,94±01,40</b>	<b>91,31±02,29</b>	<b>155,8</b>	<b>1,57</b>
avec2013	HGB	89,78±02,74	90,42±02,60	89,78±02,74	89,44±02,90	547,67	00,11
IS10_pc	RF	88,71±03,95	89,43±03,82	88,71±03,95	88,27±04,24	09,00	00,35
avec2011	HGB	88,71±01,32	88,96±01,27	88,71±01,32	88,43±01,41	465,95	00,05
emobase2010	LDA	87,63±02,01	87,72±02,22	87,63±02,01	87,52±01,93	00,25	00,28
emobase	RF	87,10±02,63	88,13±01,54	87,10±02,63	86,52±03,20	08,17	00,21
IS12_st	XGB	87,10±02,28	87,63±02,75	87,10±02,28	86,87±02,13	65,27	00,66
IS12_stc	XGB	86,02±04,23	86,25±04,25	86,02±04,23	85,81±04,33	68,73	00,58
IS11_ss	HGB	86,02±01,52	86,48±01,00	86,02±01,52	85,99±01,33	1054,70	00,50
emo_large	HGB	85,48±03,95	86,00±04,13	85,48±03,95	85,14±03,86	1637,08	00,17
IS13_ComParE	HGB	85,48±02,63	86,80±01,31	85,48±02,63	84,86±03,41	1546,62	00,70
IS09_emotion	RF	84,41±03,04	85,80±03,28	84,41±03,04	83,42±03,31	07,37	00,30
IS10_paraling	ETrees	84,41±02,74	85,93±03,94	84,41±02,74	83,58±02,55	05,67	00,42

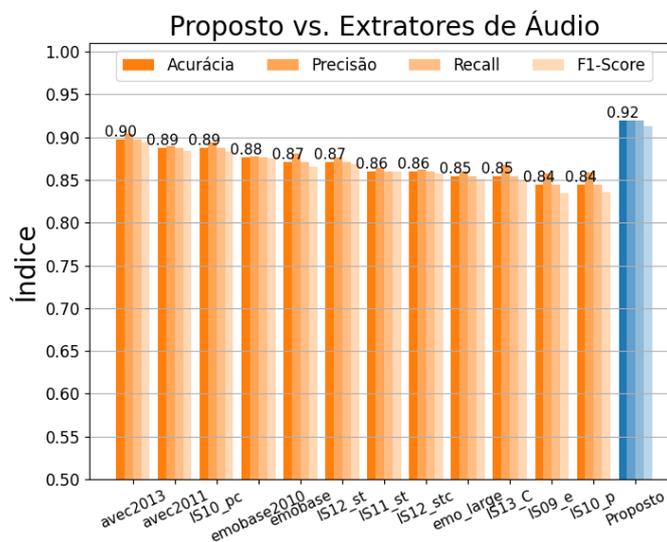
Fonte: Elaborada pelo autor.

O extrator IS12\_stc combinado com o modelo XGB apresentou 87% com maior variação, apresentando o maior desvio padrão de 4,23. Pode ser considerado o extrator com menor estabilidade quanto as taxas de acerto.

De outras formas, os extratores IS10\_paraling e IS09\_emotion apresentam os piores resultados em termos de taxas de acerto, alcançando taxas de acerto abaixo de 85%. Por outro lado, os extratores AVEC das versões de desafio 2011 e 2013 (Audio/Visual Emotion Challenge)

obtiveram os resultados mais expressivos, com acurácia de 89% e 90%.

Figura 38 – Resultados em relação aos extratores de recursos de áudio.



Fonte: Elaborada pelo autor.

A Figura 38 apresenta resultados proeminentes da metodologia proposta, com taxas de sucesso de aproximadamente 92%, superiores a todos os outros métodos em cerca de 2%. Dito isto, é possível concluir que os demais extratores tiveram dificuldades mais significativas na representação das amostras quando comparados à metodologia proposta. Acredita-se que a análise do sinal sob a perspectiva dos componentes espectrais seja menos sensível a ruídos e variações de voz em amostras da mesma classe.

Por fim, a Tabela 8 compara com trabalhos recentes apresentados na literatura especializada e tem como objetivo classificar pacientes saudáveis e HD, em todos os casos, utilizando sinais de voz. Nota-se que o trabalho de Riad *et al.* (2020) apresenta taxas de sucesso 35% inferiores à abordagem proposta. Por outro lado, o trabalho de Perez *et al.* (2018) apresenta resultados próximos a 90%, utilizando uma rede neural recorrente para analisar o sinal do paciente. Em todos os casos, a metodologia proposta apresentou resultados superiores quanto as taxas de acerto utilizando sinais de voz, indicando que a metodologia é promissora para o auxílio ao diagnóstico

## 5.6 Conclusões do capítulo

Neste Capítulo foi proposta uma aplicação da metodologia apresentada nesta tese em um cenário monocanal, utilizando como insumo para o algoritmo de extração proposto sinais de

Tabela 8 – Comparativamente com abordagens da literatura recente, os trabalhos abaixo visam classificar pacientes com doença de Huntington utilizando sinais de voz.

Autor	Abordagem	Desempenho
<b>Proposta</b>	<b>Grafo direcionado sobre componentes espectrais com ResNet-18</b>	<b>91,94% Acc, 92% Precisão, 91,94% Recall, 91,31% F1-Score</b>
Riad <i>et al.</i> (2020)	Regressão logística	56% Acc, 46% F1-Score
Perez <i>et al.</i> (2018)	k-NN, DNN, LSTM-RNN	87% Acc, 87% F1-Score
Vásquez-Correa <i>et al.</i> (2021)	CNNs	85,2% Acc, 85,5% Esp., 86,1% Sens., 94% AUC
Romana <i>et al.</i> (2020)	Filtered Vowel Distortion Measures (FVDM) with Logistic regression	80% Acc
Subramanian <i>et al.</i> (2023)	RF, SVM, Regressão logística	73% Acc, 76% Sen., 69% Esp., 72% AUC

Fonte: Elaborada pelo autor.

voz, registrados durante a leitura do dois poemas. O conjunto de amostras utilizado representou um problema de classificação binário, contendo amostras de pacientes saudáveis e amostras de pacientes com a doença de Huntington, classificada como doença neurodegenerativa pela literatura médica.

Os resultados indicam que a proposta do marcador espectral contribui para que o grafo direcionado, representado por diferentes matrizes combinadas em um tensor espectral, apresentasse resultados que superam extratores de características clássicos como wavelets, com uma diferença de até 24% de Acurácia, quando comparado ao Wavelet SYM8. Além de superar extratores clássicos baseados no domínio da frequência, como componentes de magnitude Cepstral e Fourier. Quando combinado com o modelo de classificação ResNet-18, as métricas do método proposto apresentam resultados promissores de Acurácia (91,94%), Precisão (92%) e Recall (91,94%), tais índices mostraram-se superiores a todos aqueles alcançados pelas abordagens clássicas.

A metodologia proposta, portanto, mostrou índices que superam os extratores especializados no processamento de voz. Finalmente, concluímos que a aplicação proposta neste capítulo apresenta resultados bastante promissores, podendo ser útil para o avanço nas aplicações clínicas do diagnóstico da doença de Huntington usando sinais de voz.

## 6 RESULTADOS EM CONTEXTO MULTICANAL: CLASSIFICAÇÃO DA DOENÇA DE ALZHEIMER COM MÚLTIPLOS CANAIS DE EEG.

Neste capítulo, exploramos a aplicação do descritor de sinais em amostras de pacientes saudáveis e diagnosticados com a Doença de Alzheimer, especialmente considerando a classificação multicanal com sinais de Eletroencefalograma (EEG).

### 6.1 A doença de Alzheimer

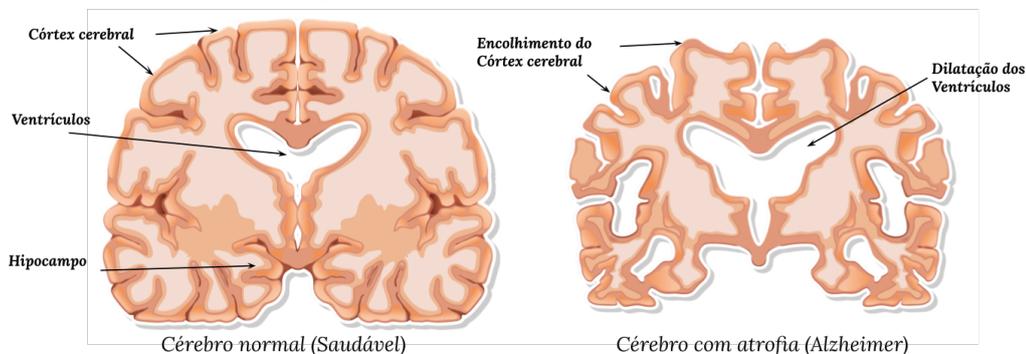
A doença de Alzheimer (AD, do inglês *Alzheimer's Disease*) é uma das enfermidades neurodegenerativa mais conhecidas sendo caracterizada por um declínio progressivo nas funções cognitivas, memória e habilidades de raciocínio. A identificação precoce dessa patologia é crucial para o tratamento eficaz e a gestão do recondicionamento clínico do paciente (PARIHAR; SWAMI, 2023). A doença de Alzheimer, conforme Hosseini *et al.* (2020), muitas vezes observada por EEG, também possui características desafiadoras quando se trata da interpretação desse sinal.

Segundo Hosseini *et al.* (2020), dentre os desafios no diagnóstico da doença via sinais de EEG, ressaltam-se a variabilidade interindividual, uma vez que trata-se de uma doença neurodegenerativa com uma apresentação variável entre os pacientes. Isso resulta em uma ampla gama de padrões de EEG, além da superposição com outras doenças, já que os padrões de EEG associados à AD podem sobrepor-se a outras doenças, tornando difícil diferenciar com precisão a AD de outras condições, como demência vascular ou comprometimento cognitivo leve.

Dentre os principais danos cerebrais apresentados por pacientes com esta patologia está a atrofia cerebral que ocorre em decorrência da morte progressiva de células nervosas. O acúmulo anormal de proteínas, como as placas de beta-amiloide e os emaranhados neurofibrilares na doença de Alzheimer podem interferir nas funções celulares normais e levar à morte neuronal. Esta condição afeta regiões responsáveis pela memória, atrofia do Hipocampo, e capacidade de aprendizado e raciocínio, atrofia do córtex cerebral (HOSSEINI *et al.*, 2020).

Estudos recentes têm explorado diversas abordagens de aprendizado profundo e modelos de inteligência artificial para aprimorar a interpretação dos sinais de EEG e a detecção de padrões associados ao Alzheimer (BATTAGLIA *et al.*, 2023; SIBILANO *et al.*, 2023; LU *et al.*, 2023; KHALIFA *et al.*, 2023; PARRA *et al.*, 2023). Neste sentido, em particular, a utilização de redes neurais profundas em sinais de EEG, registrados em estado de repouso, tem mostrado uma capacidade significativa de classificar estágios precursor ou avançados da doença. Tais avanços representam passos importantes para o diagnóstico precoce que contribui para a

Figura 39 – Atrofia cerebral devido ao acúmulo anormal de proteínas, como as placas de beta-amiloide e os emaranhados neurofibrilares na doença de Alzheimer e afeta regiões responsáveis pela memória, atrofia do Hipocampo, e capacidade de aprendizado e raciocínio, atrofia do córtex cerebral. Adaptado de Hosseini *et al.* (2020).



Fonte: Elaborada pelo autor.

eficácia da intervenção terapêutica e melhoria da sobrevivência do paciente (SIBILANO *et al.*, 2023; KHALIFA *et al.*, 2023).

Neste trabalho, com foco em avançar no campo do diagnóstico de Alzheimer, é proposta a aplicação do método de extração de atributos visando a descrição do sinal de EEG e toma como base a construção de grafos direcionados para a análise espectral dos sinais dos eletrodos. Esta aplicação sugere a identificação dos marcadores espectrais como indícios para a detecção de doenças neurodegenerativas com base no comportamento do sinal no domínio da frequência. Essa metodologia não só complementa, mas também sugere um mecanismo descritor dos sinais, visando assim fomentar a pesquisa e o diagnóstico clínico da doença de Alzheimer.

## 6.2 Trabalhos relacionados a detecção da doença

No capítulo inicial desta tese, especificamente na Seção 1.3.2.1, apresentamos diferentes abordagens para extração de atributos em análise envolvendo múltiplos canais de sinais que contribuem para a pesquisa no âmbito deste capítulo além de apresentar os trabalhos relacionados que são utilizados com fins comparativos nos resultados desta aplicação.

## 6.3 Amostras de pacientes

Nas seções a seguir, são apresentados cinco diferentes conjuntos de dados, produzidos a partir de sinais de EEG, considerando pacientes saudáveis, com a doença de Alzheimer e outras patologias. Em cada caso, são explorados os detalhes de aquisição e número de amostras.

### 6.3.1 Amostras de pacientes: Doença de Alzheimer vs. Saudável

#### 6.3.1.1 Dataset I

O conjunto de dados EEG utilizado nesta pesquisa foi disponibilizado por pesquisadores da Universidade Estadual da Flórida (VICCHIETTI *et al.*, 2023). Os sinais de EEG foram categorizados em quatro grupos distintos da seguinte forma: 12 indivíduos idosos em boa saúde, com os olhos abertos e focados visualmente (conjunto  $A_1$ ); 12 participantes idosos saudáveis com os olhos fechados (conjunto  $B_1$ ). Totalizando 24 pacientes saudáveis (conjunto  $A_1B_1$ ); 80 indivíduos com doença de Alzheimer (conjunto  $C_1$ ), também com os olhos abertos e focados visualmente; e mais 80 pacientes de AD (conjunto  $D_1$ ), desta vez com os olhos fechados. Os 160 pacientes de AD (conjunto  $C_1D_1$ ) foram diagnosticados com base nos critérios delineados no Instituto Nacional de Distúrbios Neurológicos e Acidentes Vasculares Cerebrais (NINCDS) e na Associação de Distúrbios Relacionados à Doença de Alzheimer (ADRDA). As gravações de EEG consistiram em segmentos de 8 segundos, com uma frequência de amostragem de 128 Hz, coletados usando 19 canais (Fp1, Fp2, F3, F4, F7, F8, Fz, C3, C4, Cz, P3, P4, Pz, T3, T4, T5, T6, O1, O2) seguindo o sistema internacional 10–20 para posicionamento de eletrodos. Esses sinais foram previamente filtrados para restringir sua faixa de frequência para 0,5–30 Hz.

#### 6.3.1.2 Dataset II

Este conjunto de amostras foi disponibilizado pelos trabalhos de Pineda *et al.* (2020), os dados de EEG foram registrados em 19 diferentes locais do crânio (Fp1, Fp2, Fz, F3, F4, F7, F8, Cz, C3, C4, T3, T4, Pz, P3, P4, T5, T6, O1 e O2) e seguem o sistema internacional 10-20 de posicionamento de eletrodos. Os registros abrangem quatro grupos distintos: Grupos com os olhos abertos enquanto se concentram visualmente (conjunto  $A_2$ ) e com os olhos fechados (conjunto  $B_2$ ), que representam controles saudáveis com 24 indivíduos idosos saudáveis (conjunto  $A_2B_2$ ), todos livres de quaisquer distúrbios neurológicos ou psiquiátricos. Além disso, o grupo C representa pacientes analisados com os olhos abertos (conjunto  $C_2$ ), e o grupo D com os olhos fechados (conjunto  $D_2$ ), compreendem 24 pacientes com doença de Alzheimer (conjunto  $C_2D_2$ ). Diagnosticados de acordo com os critérios do NINCDS e do ADRDA. Os segmentos de EEG, com duração de 8 segundos e filtrados na faixa de 1 a 30 Hz, foram registrados a uma frequência de amostragem de 128 Hz usando a estação de trabalho *Biologic Systems Brain Atlas III Plus*.

### 6.3.2 Amostras de pacientes: Doença de Alzheimer vs. Outras

#### 6.3.2.1 Dataset III

Este conjunto de dados engloba registros de EEG durante o estado de repouso com os olhos fechados de um total de 88 sujeitos. Dentre eles, 36 foram diagnosticados com doença de Alzheimer (conjunto  $AD_1$ ), 23 com Demência Frontotemporal (conjunto FTD), e 29 eram saudáveis (conjunto  $HC_1$ ). O estado cognitivo e neuropsicológico foi avaliado utilizando o Mini-Exame do Estado Mental Internacional (MMSE). Os registros foram realizados no 2º Departamento de Neurologia do Hospital Geral (AHEPA) de Tessalônica, na Grécia. Foi utilizado um dispositivo clínico de EEG Nihon Kohden EEG 2100, equipado com 19 eletrodos no couro cabeludo (Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1 e O2) com base no sistema internacional 10-20. Cada sessão seguiu o protocolo clínico, com os participantes sentados e os olhos fechados. Este conjunto foi disponibilizado e descrito por Miltiadous *et al.* (2023b).

#### 6.3.2.2 Dataset IV

Em 2012 e 2013, o Centro Neurolesi Bonino Pulejo, na Itália, recrutou 109 participantes. Este grupo incluiu 86 pacientes — 37 homens e 49 mulheres — e 23 amostras de controle saudáveis ( $HC_2$ ), com 13 homens e 10 mulheres. O conjunto contém amostras de pacientes com a doença de Alzheimer e Comprometimento Cognitivo Leve (MCI, do inglês *Mild Cognitive Impairment*) e segue os padrões da OMS. Os sinais de EEG foram obtidos usando 19 eletrodos posicionados de acordo com o Sistema Internacional 10-20. De cada sinal, foram selecionados os 180 segundos centrais (ou seja, de 60 a 240 segundos) para eliminar artefatos iniciais e finais na gravação do EEG. Além disso, cada sinal foi convertido para 256 Hz para padronizar a frequência de amostragem. Este conjunto de dados foi previamente avaliado por especialistas, sendo descrito e disponibilizado por (FISCON *et al.*, 2018).

#### 6.3.2.3 Dataset V

Este conjunto de dados foi compilado a partir de estudos anteriores, todos utilizando o sistema de EEG sem fio Stat X24 da Advanced Brain Monitoring, Carlsbad, EUA. O Stat

X24 é um sistema de EEG leve, alimentado por bateria e aprovado pela FDA<sup>1</sup>, que registra 20 canais de EEG seguindo os padrões internacionais 10-20. Neste conjunto, um total de 124 participantes saudáveis do grupo de controle e 89 pacientes diagnosticados com várias doenças neurodegenerativas, incluindo Comprometimento Cognitivo Leve (MCI) e doença de Alzheimer (AD). A aquisição de EEG envolveu a gravação da atividade de EEG basal durante o estado de repouso, com ambos os olhos abertos e fechados. Os sinais abrangem registros de cinco minutos com os olhos abertos e fechados, coletados e subsequentemente filtrados passa-banda na faixa de 1-40 Hz. Este conjunto de dados foi previamente avaliado por especialistas e é descrito no artigo de Meghdadi *et al.* (2019).

Tabela 9 – Resumo dos *datasets* I e II (Doença de Alzheimer vs. Saudável).

Dataset	Condição dos pacientes	Amostras	Olho fechado
Dataset I	Controle Saudáveis (A <sub>1</sub> )	12	Não
	Controle Saudáveis (B <sub>1</sub> )	12	Sim
	Doença de Alzheimer (C <sub>1</sub> )	80	Não
	Doença de Alzheimer (D <sub>1</sub> )	80	Sim
Dataset II	Controle Saudáveis (A <sub>2</sub> )	12	Não
	Controle Saudáveis (B <sub>2</sub> )	12	Sim
	Doença de Alzheimer (C <sub>2</sub> )	12	Não
	Doença de Alzheimer (D <sub>2</sub> )	12	Sim

Fonte: Elaborada pelo autor.

Tabela 10 – Resumo dos *datasets* III, IV e V (Doença de Alzheimer vs. Outras patologias).

Dataset	Condição dos pacientes	Amostras	Olho fechado
Dataset III	Doença de Alzheimer (AD <sub>1</sub> )	36	Sim
	Demência Frontotemporal (FTD)	23	Sim
	Controle Saudáveis (HC <sub>1</sub> )	29	Sim
Dataset IV	Doença de Alzheimer (AD <sub>2</sub> )	49	Sim
	Comprometimento Cognitivo Leve (MCI <sub>1</sub> )	37	Sim
	Controle Saudáveis (HC <sub>2</sub> )	23	Sim
Dataset V	Doença de Alzheimer (AD <sub>3</sub> )	13	Não
	Comprometimento Cognitivo Leve (MCI <sub>2</sub> )	18	Não
	Controle Saudáveis (HC <sub>3</sub> )	62	Não
	Doença de Alzheimer (AD <sub>4</sub> )	9	Sim
	Comprometimento Cognitivo Leve (MCI <sub>3</sub> )	19	Sim
	HControle Saudáveis (HC <sub>4</sub> )	60	Sim

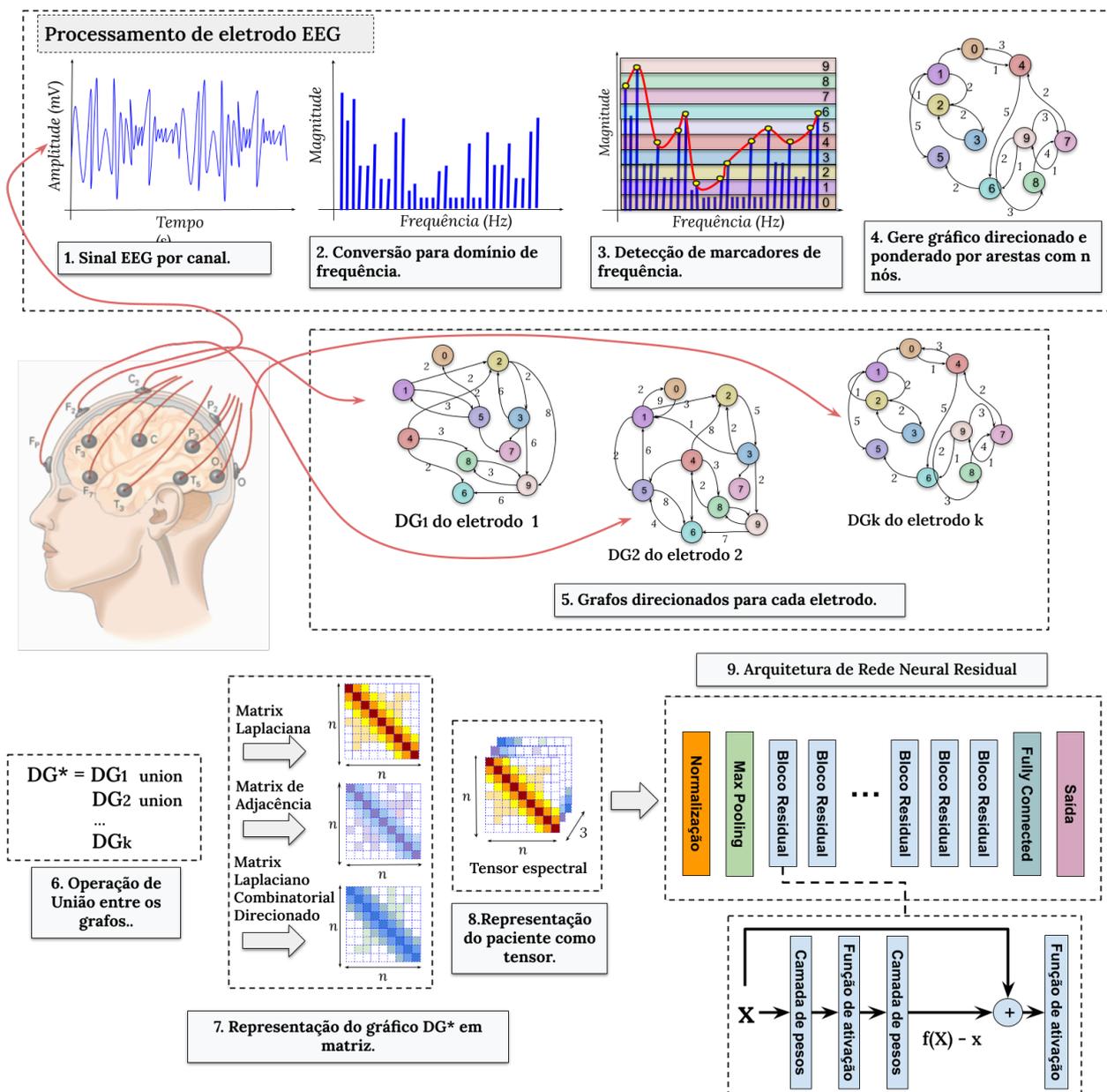
Fonte: Elaborada pelo autor.

Por fim, a Tabela 9 sumariza os conjuntos de dados com amostras das classes de Alzheimer em relação aos saudáveis. Por outro lado, a Tabela 10 apresenta o resumo dos conjuntos de dados com amostras de Alzheimer e outras patologias neurodegenerativas.

<sup>1</sup> A Food and Drug Administration (FDA) é uma agência federal do Departamento de Saúde e Serviços Humanos dos Estados Unidos

## 6.4 Metodologia

Figura 40 – Metodologia para extração dos diferentes sinais de eletroencefalograma. Para cada eletrodo, um grafo direcionado de arestas ponderadas é desenvolvido e representado como tensor para classificação via rede neural residual.



Fonte: Elaborada pelo autor.

Conforme ilustrado na Figura 40 que ilustra a metodologia desta aplicação, uma vez que o sinal de EEG foi registrado, é convertido para o domínio da frequência, via DFT (Passo 1 e 2). Em seguida, com base no tipo de marcador espectral, o sinal é analisado na busca dos

coeficientes de magnitude. Neste sentido, um coeficiente de magnitude pertence ao intervalo de magnitude espectral  $i$  se a  $k$ -ésima componente da DFT, representada por  $\phi(k)$ , apresenta a seguinte propriedade  $P$ :

$$P = \{k \in i \mid l_{inf}(i) < \phi(k) \leq l_{sup}(i)\} \quad (6.1)$$

Cada transição de um marcador no nível  $i$  para o nível  $j$  representa uma conexão entre o nó  $n_i$  e o nó  $n_j$ . Vale ressaltar que apenas os pontos marcadores serão considerados para a criação das arestas. Após identificar todas as transições dos marcadores, o mapa de transição é construído considerando a transição  $t = (n_i, n_j)$ . O número de repetições da transição  $t$  define o valor de  $w_{ij}$ , representando o peso da aresta (Passos 3 a 4).

O procedimento descrito anteriormente é aplicado a cada sinal de EEG registrado no paciente. Deste modo, temos até este ponto um grafo direcionado, ou digrafo, com arestas ponderadas, para cada eletrodo (Passo 5). Neste cenário, temos então uma situação que avalia múltiplos sinais, contexto conhecido como multicanal. Dito isto, como forme de consolidar as estruturas que representam os sinais de cada eletrodo, a metodologia sugere a operação de união, envolvendo cada grafo e produzindo um grafo que representa as características globais, o grafo  $DG^*$  (Passo 6). Seja  $DG_i$  o grafo correspondente ao sinal do eletrodo  $i$  e  $k$  indicando o número de eletrodos, temos então que  $DG^*$  é construído a partir da operação:

$$DG^* = \bigcup_{i=1}^k DG_i \quad (6.2)$$

Com base no grafo  $DG^*$ , são produzidas as matrizes de adjacência (A), que determina quais vértices estão diretamente conectados, e a existência de arestas entre os vértices é fundamental para entender a topologia geral das arestas; A matriz Laplaciana (L) que representa a relação de conectividade entre os vértices e as arestas direcionadas, destacando componentes fortemente conectados no gráfico DG; e por fim, a matriz Laplaciana Combinatória (CL), que apresenta diretamente a estrutura do grafo em termos de conectividade sem enfatizar a direção entre os vértices (Passo 7).

Finalmente, o tensor de três dimensões ( $N \times N \times 3$ ), representando as três matrizes do digrafo com  $N$  vértice, consolida as matrizes. Cada matriz representa uma perspectiva diferente de análise do digrafo  $DG^*$ . Temos então um único tensor espectral representando a amostra do paciente (Passo 8). Deste ponto em diante, segue-se para a fase de avaliação de modelos de classificação dos tensores, como tarefa de classificação das amostras em saudáveis e patológicas.

### 6.4.1 Avaliação de Desempenho

Para avaliar o desempenho da metodologia quanto a tarefa de classificação das amostras, esta metodologia considera as seguintes métricas: Precisão (Acc), Precisão (Pre), Especificidade (Esp), AUC, Recall (Rec) e o F1-Score. Por outro lado, sob o ponto de vista de separabilidade e potencial de representação das classes, avalia-se as amostras sob o ponto de vista do teste estatístico ANOVA, afim de verificar a rejeição ou aceitação da hipótese nula. Todas as métricas foram descritas em mais detalhes no Capítulo 2.

## 6.5 Resultados e discussão

Esta metodologia de avaliação considera diferentes arquiteturas de redes neurais profundas, que utilizam como entrada tensores tridimensionais, apresentadas no Capítulo 2. A fase de classificação das amostras geradas pela metodologia de extração considerou a estratégia de validação cruzada *Stratified K-Folds*, com  $K=5$ , esta metodologia faz amostragens estratificada das classes e tem sido utilizada em trabalhos recentes da literatura para a tarefa de classificação de doenças (GANSER *et al.*, 2021; GAMBOA *et al.*, 2022).

Os modelos avaliados foram treinados com os seguintes parâmetros: Cross-Entropy Cost Function, AdaGrad, taxa de aprendizagem igual a 0,0001 com 50 épocas. O ambiente computacional possui as seguintes configurações: Intel Core i7 2.2GHz, RAM 16GB, RTX 3050 Mobile, Python 3.7, Torch 1.10, Scikit-learn 1.0.2

Com base na análise dos componentes espectrais dos sinais de EEG, o número de vértices utilizados para construir o grafo direcionado foi avaliado considerando a estratégia de busca em grade, onde o número de níveis de magnitude é  $L = \{x \in \mathbb{N} \mid 3 \leq x \leq 30 \text{ e } x \bmod 3 = 0\}$ , ou seja, entre 3 e 30, com um passo de 3. Cada valor de  $L$  foi avaliado de acordo com o valor de acurácia apresentado pela fase de classificação dos tensores espectrais, conforme algoritmo para a seleção de parâmetros descrito no Capítulo 3. Os resultados apresentados nesta seção, consideram esta configuração para a construção do digrafo para a representação das amostras.

### 6.5.1 Teste estatístico ANOVA

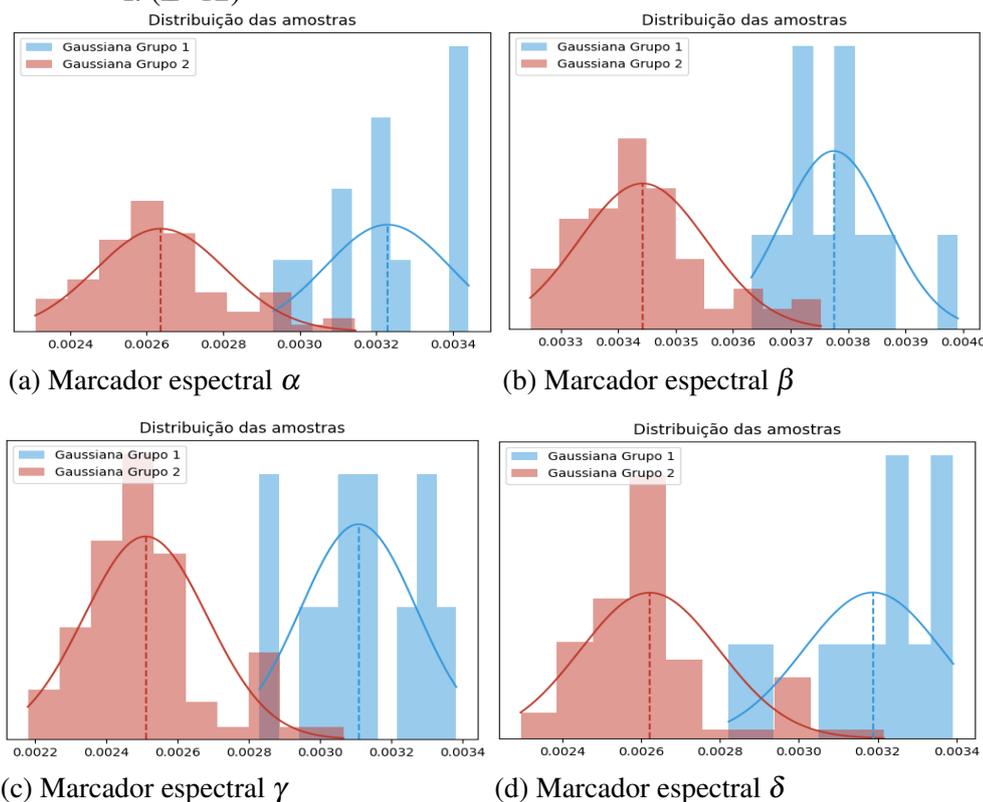
Nas seções a seguir, avaliamos as amostras extraídas em dois diferentes contextos. O primeiro cenário diz respeito a situação de um problema de classificação binário, isto é, apenas duas classes são consideradas, Seção 6.5.2. Por outro lado, representando o cenário para

classificação multiclasse, a Seção 6.5.3 apresenta os resultados para o cenário com três classes, sendo duas classes patológicas e uma representando o grupo saudável.

### 6.5.2 Teste com duas classes

A Figura 41 apresenta a distribuição das amostras após a extração com a metodologia proposta, contemplando quatro diferentes situações de acordo com os marcadores espectrais. Os resultados apresentados referem-se a extração no *dataset* I e considera as classes AD vs. HC. Com base na distribuição apresentada no histograma, é possível observar na marcação das curvas gaussianas que a projeção da média dos grupos apresenta uma diferença bastante expressiva entre as classes. O marcador espectral  $\beta$ , por exemplo, exibe dois grupos, com médias significativamente distintas, Figura 41-b.

Figura 41 – Histograma das amostras extraídas para as classes AD vs. HC do *dataset* I. (L=12)

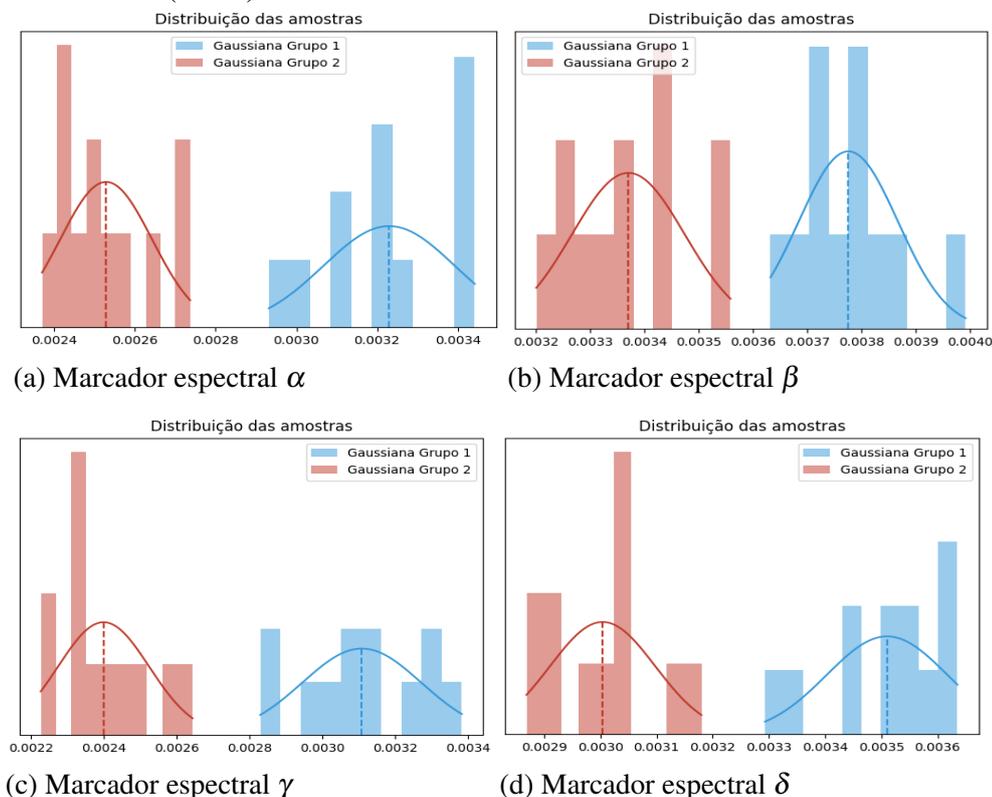


Fonte: Elaborada pelo autor.

Este comportamento além de sugerir uma distribuição com perfil gaussiano, sugere também que o método de extração consegue agrupar as amostras das duas classes em dois diferentes grupos, como fica evidente na distância das médias dos grupos. De modo similar, a

Figura 42, exibe as curvas gaussianas para amostras do *dataset* II, e também considera as classes AD vs. HC. De modo semelhante, exibe uma distribuição de grupos distintos.

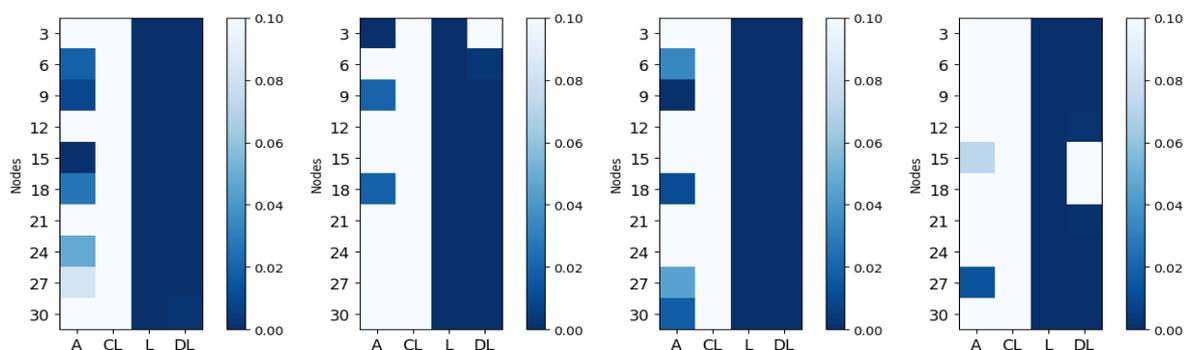
Figura 42 – Histograma das amostras extraídas para as classes AD vs. HC do *dataset* II. ( $L=12$ )



Fonte: Elaborada pelo autor.

Por outro lado, tomando como base a distribuição apresentada pelas amostras extraídas, é possível a aplicação do teste estatístico ANOVA. A Figura 43 apresenta o valor- $p$  para as amostras extraídas das classes AD vs. HC, do *dataset* I. São apresentados cenários conforme o tipo de marcador:  $\alpha$  (43-a),  $\beta$  (43-b),  $\gamma$  (43-c) e  $\delta$  (43-d). As linhas representam as diferentes configurações de número de vértices e as colunas, as matrizes A, CL, L e DL, que representam as diferentes perspectivas do grafo direcionado.

Figura 43 – Resultado do valor- $p$  no teste ANOVA para discriminação no *dataset I* (AD vs HC)



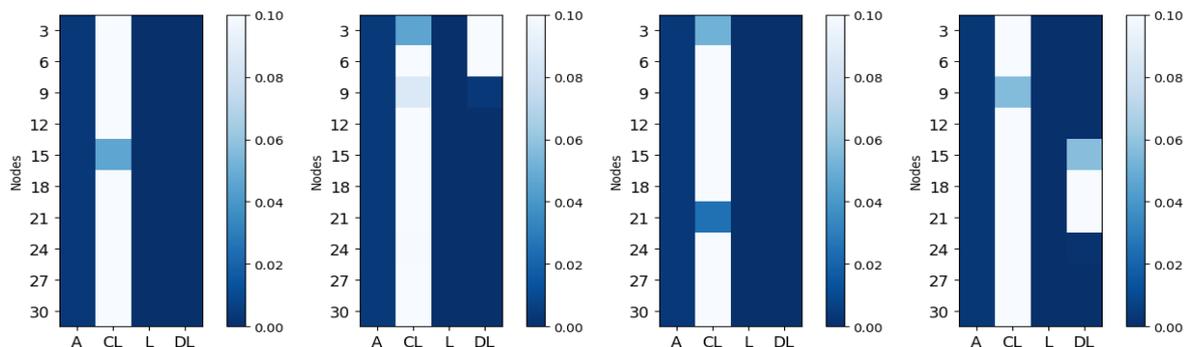
(a) Marcador espectral  $\alpha$  (b) Marcador espectral  $\beta$  (c) Marcador espectral  $\gamma$  (d) Marcador espectral  $\delta$

Fonte: Elaborada pelo autor.

É possível observar que as matrizes A, L e DL apresentam valor- $p$  inferiores ao nível de significância  $\alpha$  (0.05), como pode ser observado conforme escala ao lado. O que indica rejeição da hipótese nula e assim, indicam potencial de separação das classes. A matriz CL, entretanto, não generaliza o nível do valor- $p$  para todas as combinações de vértices e apresenta pouca relevância nos marcadores  $\alpha$  e  $\gamma$ .

De modo similar, a Figura 44 apresenta resultados relevantes para o valor- $p$  das matrizes A, L e DL, e novamente indica que estas matrizes apresentam potencial para melhor discriminar as amostras entre as classes AD vs. HC do *dataset II*. Por outro lado, a Figura 45 não apresenta valores expressivos do valor- $p$  para a maioria das combinações e marcadores para as classes AD vs HC, do *dataset III*. Com exceção da matriz DL cujo resultado indica rejeição da hipótese nula para combinações a partir de 21 vértices, nos marcadores  $\gamma$  e  $\delta$ . Para este caso, acredita-se que o procedimento de aquisição dos sinais, diferente dos *datasets I* e *II*, contempla apenas pacientes com o olho fechado durante o exame, o que pode influenciar no EEG.

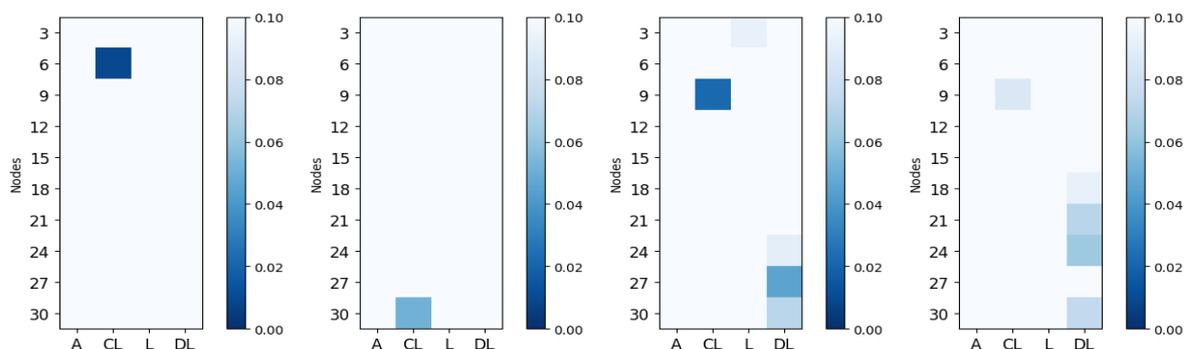
Figura 44 – Resultado do valor- $p$  no teste ANOVA para discriminação no *dataset II* (AD vs HC)



(a) Marcador espectral  $\alpha$  (b) Marcador espectral  $\beta$  (c) Marcador espectral  $\gamma$  (d) Marcador espectral  $\delta$

Fonte: Elaborada pelo autor.

Figura 45 – Resultado do valor- $p$  no teste ANOVA para discriminação no *dataset III* (AD vs HC)

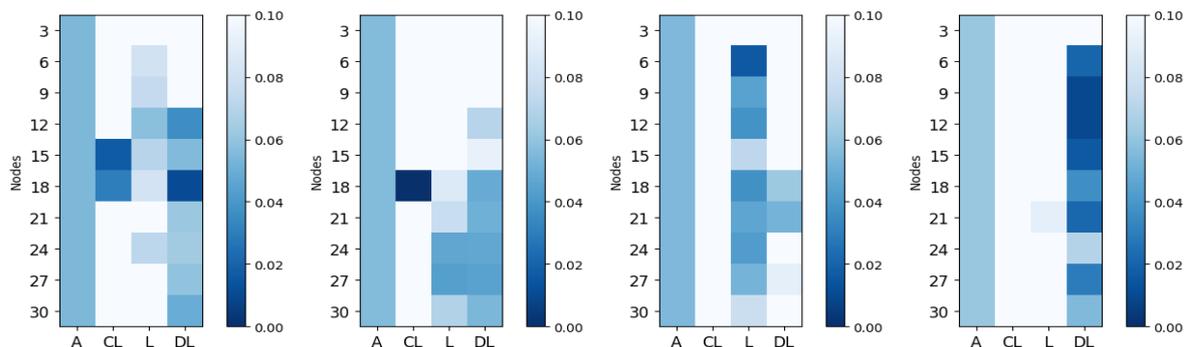


(a) Marcador espectral  $\alpha$  (b) Marcador espectral  $\beta$  (c) Marcador espectral  $\gamma$  (d) Marcador espectral  $\delta$

Fonte: Elaborada pelo autor.

Já a Figura 46 apresenta resultados para duas classes de amostras patológicas, a saber AD vs MCI. Esta combinação apresenta um desafio, já que referem-se a dois tipos de patologia com características que se sobrepõem, como apresenta Fiscon *et al.* (2018). Contudo, os resultados do valor- $p$  indicam que as matrizes A, L, DL apresentam potencial para discriminar amostras entre as classes. Ressaltando melhores resultados para os marcadores do tipo  $\alpha$  e  $\gamma$ , acima de 6 vértices.

Figura 46 – Resultado do valor- $p$  no teste ANOVA para discriminação de amostras no *dataset IV* com 2 classes (AD vs MCI)

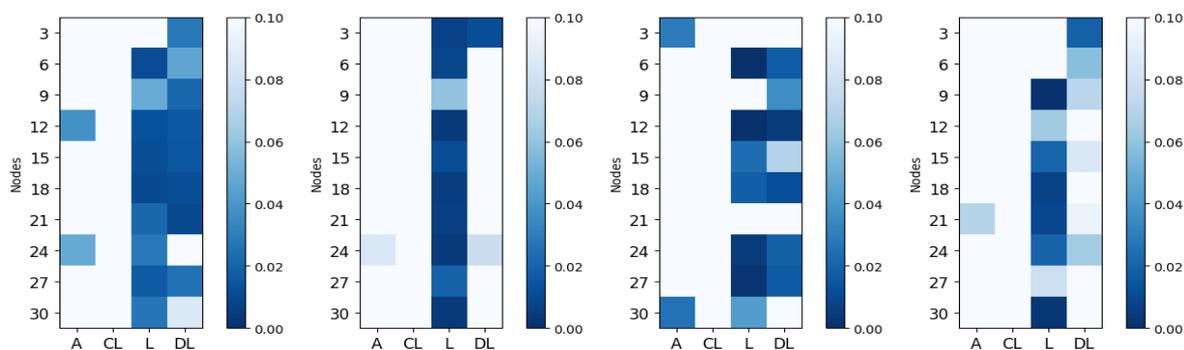


(a) Marcador espectral  $\alpha$  (b) Marcador espectral  $\beta$  (c) Marcador espectral  $\gamma$  (d) Marcador espectral  $\delta$

Fonte: Elaborada pelo autor.

Por fim, a Figura 47 apresenta o valor- $p$  para amostras das classes AD vs. HC do *dataset V*. E assim como o *dataset I* e II, os resultados indicaram relevância para as matrizes L e DL, nos marcadores  $\alpha$ ,  $\gamma$  e  $\delta$ . Estes valores indicam a importância dos picos de magnitude do sinal no domínio da frequência, o que também sugere que o descritor proposto conseguiu identificar tal propriedade e representá-la no resultado da extração.

Figura 47 – Resultado do valor- $p$  no teste ANOVA para discriminação de amostras no *dataset V* (AD vs HC)



(a) Marcador espectral  $\alpha$  (b) Marcador espectral  $\beta$  (c) Marcador espectral  $\gamma$  (d) Marcador espectral  $\delta$

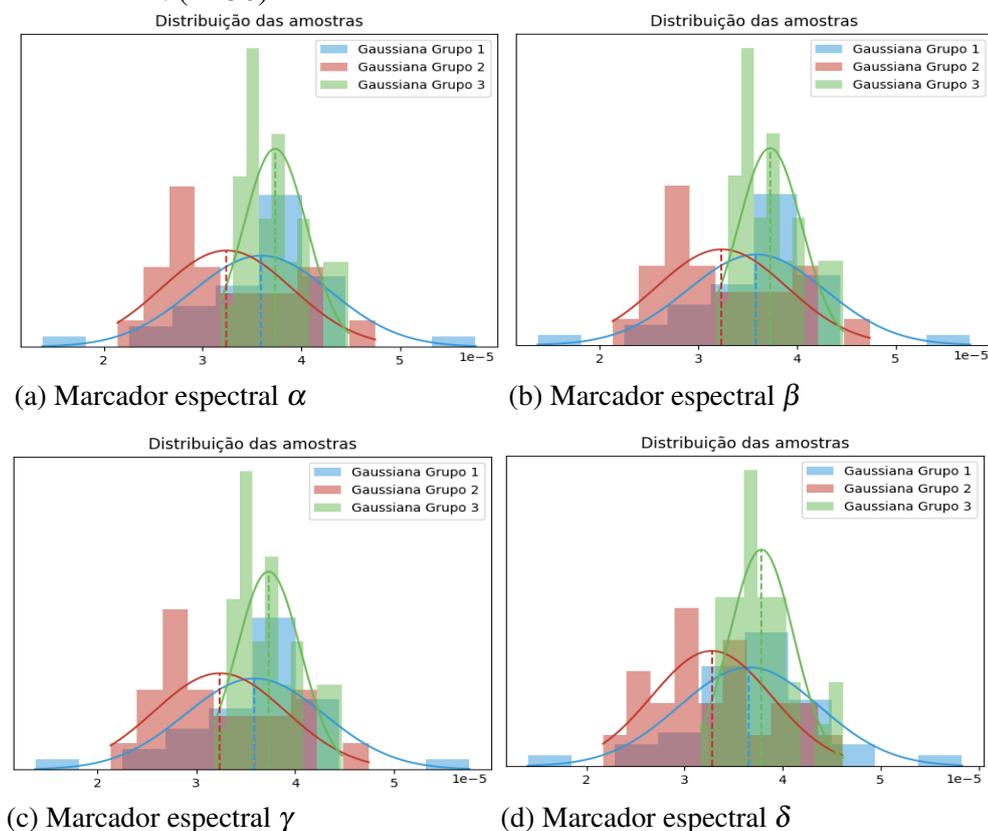
Fonte: Elaborada pelo autor.

### 6.5.3 Teste com três classes

Por outro lado, representando o cenário multiclasse, as Figuras 48 e 49 apresentam a distribuição das classes, AD (em azul), MCI (em vermelho) e HC (em verde). Nas distribuições é possível observar uma relevância para o marcador espectral  $\beta$  e  $\delta$ , em que nota-se a linha tracejada que indica a média dos grupos com um distanciamento mais relevante entre os grupos.

Para o cenário AD vs. FTD vs. HC, todos os marcadores apresentaram diferença nas médias. O

Figura 48 – Histograma da extração para as classes AD vs. FTD vs. HC do *dataset* III. ( $L=30$ )

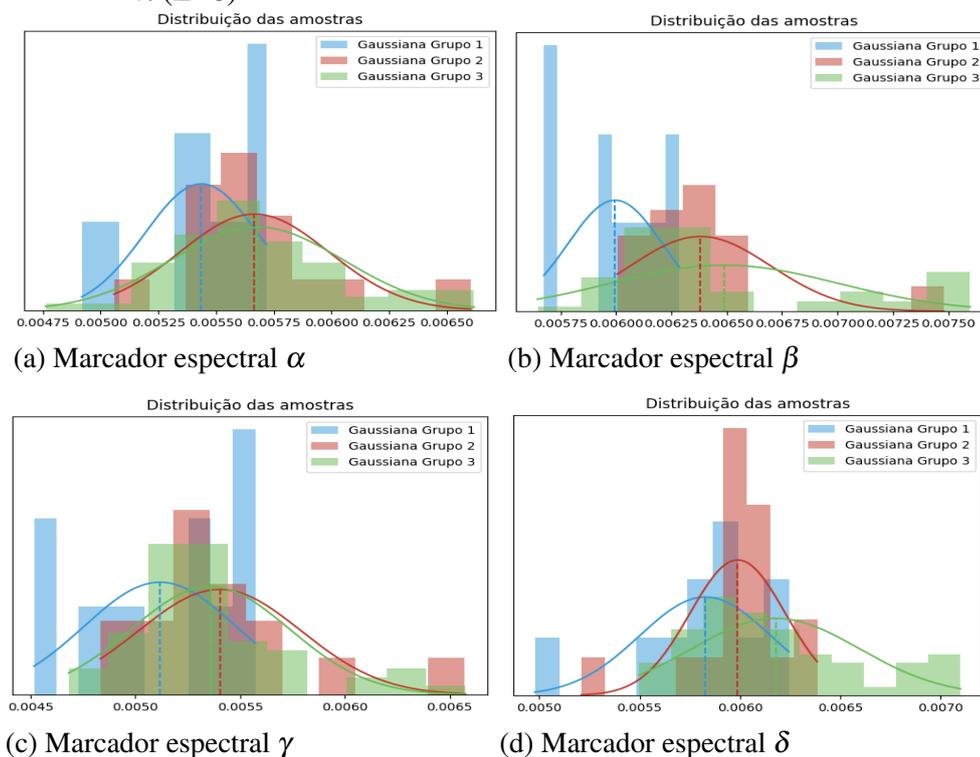


Fonte: Elaborada pelo autor.

mesmo resultado ocorre, no cenário AD vs. MCI vs. HC, com exceção apenas do marcador  $\alpha$  que apresentam uma sobreposição das classes MCI (em vermelho) e HC (em verde), indicando que a análise dos máximos locais, para este conjunto, não mostrou relevância na separação em grupos, Figura 49.

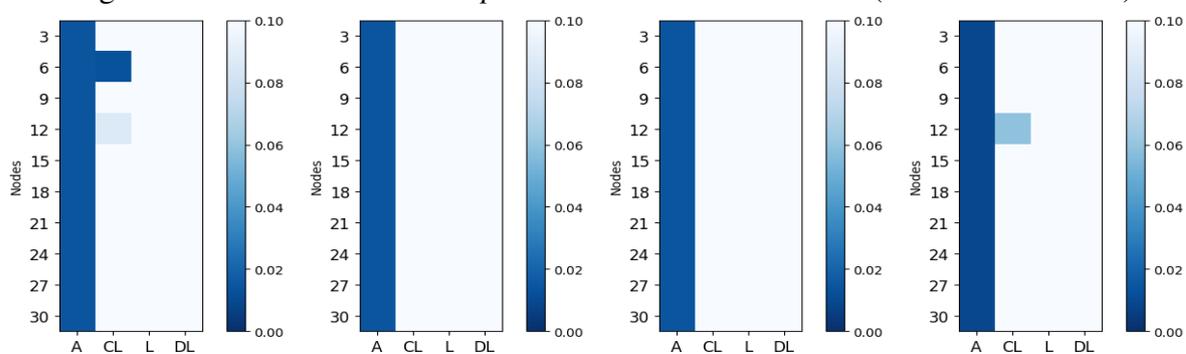
Por outro lado, o teste ANOVA para o cenário AD vs FTD vs HC, Figura 50 mostrou que apenas a matriz A evidenciou maior relevância quanto ao potencial de separação das classes, com valores- $p$  que indicam a rejeição da hipótese nula. Já para o cenário AD vs MCI vs HC, Figura 51 as matrizes L e DL apresentaram maior relevância, com destaque para a matriz L que apresentou valor- $p$  significativo na maioria das combinações de vértices.

Figura 49 – Histograma da extração para as classes AD vs. MCI vs. HC do *dataset* V. ( $L=6$ )

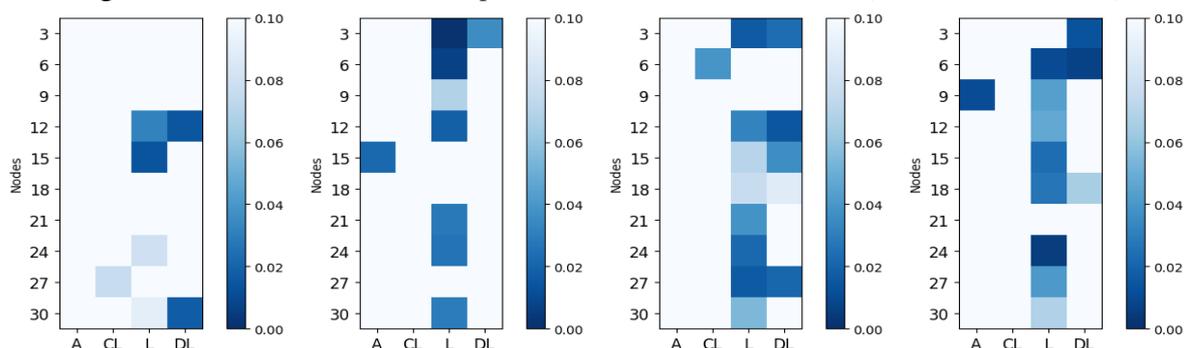


Fonte: Elaborada pelo autor.

Figura 50 – Resultado do valor- $p$  no *dataset* III com 3 classes (AD vs FTD vs HC)



Fonte: Elaborada pelo autor.

Figura 51 – Resultado do valor- $p$  no *dataset V* com 3 classes (AD vs MCI vs HC)(a) Marcador espectral  $\alpha$  (b) Marcador espectral  $\beta$  (c) Marcador espectral  $\gamma$  (d) Marcador espectral  $\delta$ 

Fonte: Elaborada pelo autor.

Diante dos resultados quanto ao teste ANOVA, é possível inferir que as matrizes A, L e DL apresentaram potencial para discriminação entre as classes, com resultados significativos para o cenário AD vs. HC. Já para os cenários multiclasse, a matriz A apresentou maior relevância para a distinção entre grupos contendo amostras da patologia FTD. Em todos os casos analisados, foi possível notar que as três matrizes, em diferentes situações, apresentam potencial discriminatório das populações de amostras.

## 6.6 Combinação dos marcadores das componentes espectrais com modelos de classificação

Os resultados são apresentados sob duas perspectivas, a primeira comparando a classificação de pacientes com Alzheimer em relação aos Saudáveis (Tabela 11) e a segunda comparando amostras de Alzheimer com outras doenças degenerativas (Tabelas 12, 13 e 14). De acordo com a Tabela 11, a combinação do classificador com a metodologia de extração proposta apresenta taxas de acerto significativamente altas, alcançando índices promissores de Acurácia (100%), Precisão (100%) e AUC (100%), para as amostras do *dataset I*, que correspondem a amostras de pacientes com olho aberto ( $A_1$ ) e também olho aberto ( $B_1$ ).

Sobre os subconjuntos  $A_1$  vs  $C_1$ , a metodologia proposta apresenta resultados ligeiramente superiores aqueles apresentados pelas abordagens de Vicchiotti *et al.* (2023) quanto a ACC, que analisa o sinal no domínio do tempo, e Pineda *et al.* (2020) quanto a AUC no *dataset A\_2* vs  $C_2$ , que combina diferentes extratores, também no domínio do tempo. A diferença apresentada indica que a análise proposta no domínio da frequência mostrou-se ligeiramente superior, uma vez que a proposta avalia as componentes espectrais do sinal, e deste modo confere robustez a ruído estático no sinal EEG.

Tabela 11 – Resultados para o melhor desempenho do método proposto nas amostras do dataset I e II, para classificação do Alzheimer em relação a pacientes saudáveis.

Dataset	Subset	Descritor	Classificador	ACC	Pre.	Esp.	AUC	Recall	F1-Score	Treino(s)	Teste(s)	
Dataset I	$A_1+B_1$ vs $C_1+D_1$	Proposto ( $\delta$ -12 vértices)	ResNet-18	100.00	100.00	100.00	100.00	100.00	100.00	28.33	0.96	
	$A_1$ vs $C_1$	Proposto ( $\delta$ -12 vértices)	ResNet-18	<b>100.00</b>	<b>15.55</b>	<b>0.84</b>						
		Vicchietti <i>et al.</i> (2023)	SVM	98.90	100	91.70	100.00	-	-	-	-	-
	$B_1$ vs $D_1$	Proposto ( $\gamma$ -12 vértices)	ResNet-18	<b>100.00</b>	<b>100.00</b>	<b>100.00</b>	<b>100.00</b>	<b>100.00</b>	<b>100.00</b>	<b>12.48</b>	<b>0.91</b>	
	$A$ vs $D$	Vicchietti <i>et al.</i> (2023)	SVM	100	100	100	100.00	-	-	-	-	-
		Proposto ( $\alpha$ -12 vértices)	ResNet-18	100.00	100.00	100.00	100.00	100.00	100.00	100.00	12.16	0.83
Dataset II	$A_2+B_2$ vs $C_2+D_2$	Proposto ( $\delta$ -12 vértices)	ResNet-18	100.00	100.00	100.00	100.00	100.00	100.00	8.72	2.15	
	$A_2$ vs $C_2$	Proposto ( $\delta$ -12 vértices)	ResNet-18	<b>100.00</b>	<b>4.61</b>	<b>1.43</b>						
$B_2$ vs $D_2$		Pineda <i>et al.</i> (2020)	SVM	-	-	-	96.53	-	-	-	-	-
	Proposto ( $\delta$ -12 vértices)	ResNet-18	<b>100.00</b>	<b>6.06</b>	<b>0.55</b>							
$A_2$ vs $D_2$	Pineda <i>et al.</i> (2020)	SVM	-	-	-	100.00	-	-	-	-	-	
	Proposto ( $\delta$ -12 vértices)	ResNet-18	<b>100.00</b>	<b>4.61</b>	<b>1.46</b>							
$B_2$ vs $C_2$	Pineda <i>et al.</i> (2020)	SVM	-	-	-	96.53	-	-	-	-	-	
	Proposto ( $\delta$ -12 vértices)	ResNet-18	<b>100.00</b>	<b>6.12</b>	<b>0.75</b>							
		Pineda <i>et al.</i> (2020)	SVM	-	100.00	100.00	100.00	-	-	-	-	

Fonte: Elaborada pelo autor.

Observa-se também que a combinação do método de extração de atributos proposto se destaca na classificação de pacientes com Alzheimer independente do estímulo durante a aquisição do sinal EEG. Isto é, seja o sinal registrado com o olho aberto ou fechado, a abordagem apresenta resultados com ACC (100) e AUC (100) para os subconjuntos  $A_1 + B_1$  vs  $C_1 + D_1$  e  $A_2 + B_2$  vs  $C_2 + D_2$ . A principal diferença quanto as amostras é que a situação com olho aberto está mais suscetível a ruídos produzidos por estímulos visuais.

Dentre os coeficientes de magnitude, destaca-se o resultado apresentado pelo marcador espectral  $\delta$ , quando combinado o grafo direcionado de 12 vértices com a rede residual ResNet. Esta combinação mostra os melhores resultados para os *datasets* I e II. Sugere-se que o marcador  $\delta$  apresentou maior potencial para o cenário de discriminação AD vs. HC. Os resultados da Tabelas 12, no *dataset* III, novamente corroboram com esta hipótese, onde a classificação de AD vs. HC apresentam taxas de ACC (93,94%), Pre (94,39%) e AUC (92,8%) promissoras; superando a abordagem de Chen *et al.* (2023) em 6% de ACC e até 10% na abordagem de Miltiadous *et al.* (2023a) que também utiliza aprendizagem profunda.

A Tabela 12 apresenta ainda o cenário AD vs DFT vs HC, em que a metodologia proposta sugere resultados próximos de 90%, superando a proposta do Chen *et al.* (2023) em mais de 9% de precisão. Já a abordagem Chen *et al.* (2023), embora utilize componentes espectrais, faz uma combinação de características globais do sinal no domínio do tempo e este resultado sugere que a análise global no domínio do tempo não beneficia o modelo, diferentemente do método proposto neste trabalho, que avalia apenas componentes espectrais.

Já a Tabela 13 apresenta os resultados para o *dataset* IV, nesta análise, são apresentados extratores que atuam também no domínio da frequência e nota-se que há um desempenho

Tabela 12 – Resultados para o melhor desempenho do método proposto nas amostras do dataset III para classificação do Alzheimer em relação a pacientes com outras patologias.

Dataset	Subset	Descritor	Classificador	ACC	Pre.	Esp.	AUC	Recall	F1-Score	Treino (s)	Teste (s)
Dataset III	AD <sub>1</sub> vs HC <sub>1</sub>	<b>Proposto (<math>\delta</math>-24 vértices)</b>	ResNet-34	93.94	94.39	91.66	92.8	93.94	93.94	41,84	2.79
		Coerência espectral	DICE-net	83.28	87.94	87.94		79.81	84.12		
		Miltiadous <i>et al.</i> (2023a)	LightGBM	76.28	79.67	76.52		76.08	77.83		
			XGBoost	75.53	78.55	74.87		76.08	77.29		
		The joint method		87.33		85.15	88.19	84.56			
	Chen <i>et al.</i> (2023)	EEG model	85.78		83.22	85.88	81.76				
		Random forests	77.01		80.94	78.32					
	AD <sub>1</sub> vs FTD vs HC <sub>1</sub>	<b>Proposto (<math>\gamma</math>-30 vértices)</b>	ResNet-101	88.89	88.79	91.66	89.95	88.89	88,56	141,37	4.26
		The joint method		79.12		78.30	80.23	77.09			
		Chen <i>et al.</i> (2023)	EEG model	76.01		75.45	76.37	73.20			
AD <sub>1</sub> vs (FTD+HC <sub>1</sub> )	<b>Proposto (<math>\alpha</math>-18 vértices)</b>	ResNet-152	93.33	93,69	94,44	93,37	93,33	93,39	122,71	1.89	

Fonte: Elaborada pelo autor.

Tabela 13 – Resultados para o melhor desempenho do método proposto nas amostras do dataset IV para classificação do Alzheimer em relação a pacientes com outras patologias.

Dataset	Subset	Descritor	Classificador	ACC	Pre.	Esp.	AUC	Recall	F1-Score	Treino (s)	Teste (s)
Dataset IV	AD <sub>2</sub> vs HC <sub>2</sub>	<b>Proposto (<math>\delta</math>-27 vértices)</b>	ResNet-18	80,56	83,36	83,67	74,44	80,56	77,52	7,51	1.14
		Fourier Coefficients	Árvore de Decisão	72.20	71.10	59.00		72.20	71.40		
		Fiscon <i>et al.</i> (2018)									
		Wavelet Coefficients	Árvore de Decisão	83.3	83.3	78.0		83.3	83.3		
		Fiscon <i>et al.</i> (2018)									
	MCI <sub>1</sub> vs HC <sub>2</sub>	<b>Proposto (<math>\beta</math>-24 vértices)</b>	ResNet-34	86,67	86,46	94,59	84,25	86,67	85,95	5,69	0.66
		Fourier Coefficients	Árvore de Decisão	71.7	78.9	79.0		71.7	71.8		
		Fiscon <i>et al.</i> (2018)									
		Wavelet Coefficients	Árvore de Decisão	91.7	91.8	91.5		91.7	91.7		
		Fiscon <i>et al.</i> (2018)									
	MCI <sub>1</sub> vs AD <sub>2</sub>	<b>Proposto (<math>\gamma</math>-21 vértices)</b>	ResNet-50	89,66	91,11	78,37	88,16	89,66	88,68	7,80	0.41
		Fourier Coefficients	Árvore de Decisão	80.2	80.2	78.5		80.2	80.1		
Fiscon <i>et al.</i> (2018)											
	Wavelet Coefficients	Árvore de Decisão	79.1	79.3	79.1		79.1	79.1			
	Fiscon <i>et al.</i> (2018)										
(MCI <sub>1</sub> +AD <sub>2</sub> ) vs HC <sub>2</sub>	<b>Proposto (<math>\beta</math>-27 vértices)</b>	ResNet-50	94,57	95,26	99,00	86,95	94,57	93,70	9,90	0.74	
	Fourier Coefficients	Árvore de Decisão	74.70	74.00	46.30		74.70	74.70			
	Fiscon <i>et al.</i> (2018)										
	Wavelet Coefficients	Árvore de Decisão	73.4	74.70	51.50		73.40	74.00			
	Fiscon <i>et al.</i> (2018)										

Fonte: Elaborada pelo autor.

bastante competitivo entre as abordagens. A proposta deste trabalho, entretanto, supera em até 9% as combinações com Coeficientes de Fourier e Wavelets considerando o cenário AD vs. MCI, sugerindo um bom potencial de separação entre duas doenças neurodegenerativas.

Por sua vez a Tabela 14 considera também o cenário amostral entre AD e MCI, os resultados indicam taxas de acurácia superiores a 96% com AUC de 97%. Comparado com o trabalho de Meghdadi *et al.* (2019), no cenário AD vs. HC, o método proposto apresenta uma especificidade de 88% em comparação com 76%. Comparada ao MCI x HC, a metodologia proposta apresenta 88,88% de ACC e supera a proposta de Meghdadi *et al.* (2019) no mesmo conjunto. Por fim, a Tabela 15 compara brevemente os cenários avaliados nesta metodologia com os resultados alcançados em trabalhos recentes na literatura especializada.

Tabela 14 – Resultados para o melhor desempenho do método proposto nas amostras do dataset V para classificação do Alzheimer em relação a pacientes com outras patologias.

Dataset	Subset	Descritor	Classificador	ACC	Pre.	Esp.	AUC	Recall	F1-Score	Treino (s)	Teste (s)
Dataset V	AD <sub>3</sub> vs HC <sub>3</sub>	<b>Proposto (<math>\gamma</math>-24 vértices)</b>	ResNet-18	98,67	98,93			98,67	98,72	31,73	2,97
	MCI <sub>2</sub> vs AD <sub>3</sub>	<b>Proposto (<math>\gamma</math>-30 vértices)</b>	ResNet-101	96,97	97,40			96,97	96,92	10,56	7,19
	MCI <sub>2</sub> vs HC <sub>3</sub>	<b>Proposto (<math>\beta</math>-3 vértices)</b>	ResNet-18	96,30	96,72			96,30	95,30	05,98	1,44
	AD <sub>4</sub> vs HC <sub>4</sub>	<b>Proposto (<math>\delta</math>-12 vértices)</b> Meghdadi <i>et al.</i> (2019)	ResNet-18 LDA	98,61	98,67	88,88	94,44	98,61	98,49	10,93	0,42
	MCI <sub>3</sub> vs AD <sub>4</sub>	<b>Proposto (<math>\alpha</math>-6 vértices)</b>	ResNet-50	96,29	97,22	94,44	97,22	96,29	96,39	5,67	0,11
	MCI <sub>3</sub> vs HC <sub>4</sub>	<b>Proposto (<math>\delta</math>-24 vértices)</b> Meghdadi <i>et al.</i> (2019)	ResNet-101 LDA	97,53	97,75	88,88	94,44	97,53	97,34	86,64	2,76
	AD <sub>3</sub> +AD <sub>4</sub> vs HC <sub>3</sub> +HC <sub>4</sub>	<b>Proposto (<math>\delta</math>-15 vértices)</b>	ResNet-101	97,22	97,47			97,22	96,73	96,43	1,28
	AD <sub>3</sub> +AD <sub>4</sub> vs MCI <sub>2</sub> +MCI <sub>3</sub>	<b>Proposto (<math>\delta</math>-12 vértices)</b>	ResNet-50	95,00	95,02			95,00	94,93	19,96	0,51

Fonte: Elaborada pelo autor.

Tabela 15 – Sumário comparativo com trabalhos relacionados a classificação de Alzheimer.

Cenário	Estudo	Ano	Subset	Acurácia / AUC
AD vs HC	<b>Proposto</b>	<b>2023</b>	I. 80 AD,12 HC	<b>I. 100.00 / 100.00</b>
			II. 12 AD,12 HC	<b>II. 100.00 / 100.00</b>
			III. 36 AD,29 HC	<b>III. 93.94 / 92.80</b>
			IV. 49 AD,23 HC	<b>IV. 94.57 / 86.95</b>
			V. 21 AD,122 HC	<b>V. 97.22 / 90.90</b>
	Vicchiotti <i>et al.</i> (2023)	2023	80 AD, 12 HC	98.90
	Chen <i>et al.</i> (2023)	2023	24 AD,24 HC	- / 88.19
	Miltiadous <i>et al.</i> (2023a)	2023	36 AD,29 HC	83.28
	Lopes <i>et al.</i> (2023)	2023	34 AD,20 HC	87.30
	Dogan <i>et al.</i> (2023)	2022	12 AD,11 HC	92.01
Pineda <i>et al.</i> (2020)	2020	12 AD,12 HC	- / 96.53	
Meghdadi <i>et al.</i> (2019)	2019	9 AD,60 HC	- / 80.00	
AD vs MCI vs HC	<b>Proposto</b>	<b>2023</b>	9 AD,19 MCI, 60 HC	<b>97.53 / 97.53</b>
	Araújo <i>et al.</i> (2022)	2022	11 AD,8 MCI,11 HC	81.30
	Meghdadi <i>et al.</i> (2019)	2019	9 AD,19 MCI, 60 HC	- / 61.00
	Ruiz-Gómez <i>et al.</i> (2018)	2018	74 AD + MCI,37 HC	78.43
AD vs FTD vs HC	<b>Proposto</b>	<b>2023</b>	36 AD, 23 FTD, 29 HC	<b>88.89 / 89.95</b>
	Chen <i>et al.</i> (2023)	2023	36 AD, 23 FTD, 29 HC	79.12 / 80.23

Fonte: Elaborada pelo autor.

## 6.7 Conclusões do capítulo

Neste Capítulo foi proposta uma aplicação da metodologia apresentada nesta tese em um cenário multicanal, utilizando sinais de EEG como insumo para o algoritmo de extração, contemplando cinco diferentes conjuntos de amostras de pacientes reais. Nos *datasets* I, II e IV, o conjunto de amostras utilizado representou um problema de classificação binário, que contém amostras de pacientes saudáveis e amostras de pacientes com a doença de Alzheimer, classificada como doença neurodegenerativa pela literatura médica. Para estes casos, foi avaliada a classificação de amostras entre saudáveis e doentes. Por outro lado, nos *datasets* III e V, há também outras doenças neurodegenerativas além do Alzheimer, tais como Comprometimento Cognitivo Leve (MCI) e Demência Frontotemporal (FTD), estes casos representam então cenário de classificação multiclasse.

Os resultados indicam que a proposta dos marcadores espectrais sobre o sinal no domínio da frequência contribui para que o tensor espectral como representação da amostra fosse

compatível ou superasse abordagens da literatura recente. Quanto ao cenário de classificação binária, os resultados apresentaram índices de acurácia, especificidade e AUC próximos a 100%, sendo igual ou superior aos métodos que processaram os mesmos *dataset* I e II. Já em relação ao *dataset* IV, o cenário de classificação *ADvsHC*, apresentou índices próximos a 90%, ainda assim, superando metodologias que processaram o mesmo *dataset*. Por outro lado, quanto ao cenário multiclasse, a metodologia proposta apresentou acurácia e AUC de 97,53% para o cenário *ADvsMCIsHC*. Já para o outro cenário, contendo a doença FTD, os resultados foram próximos a 90%, para os mesmos índices, ainda assim, superando metodologias executadas no mesmo *dataset*.

Vale destacar ainda que a proposta de construir digrafos a partir do sinal dos eletrodos em um único tensor e representa-lo sob as diferentes perspectivas de matrizes indicou potencial promissor quando comparado o valor-*p* em todos os cinco *dataset* avaliados. Ressaltando que pelo menos umas das matrizes que compõem o tensor espectral apresentou rejeição da hipótese nula, independentemente do cenário seja ele discriminação de duas classes, tal como *ADvs.HC*, ou no cenário multiclasse, discriminando AD de saudáveis e outras patologias. Indicando assim que a metodologia pode ser relevante no auxílio a tomada de decisão do especialista quanto ao diagnóstico da doença de Alzheimer.

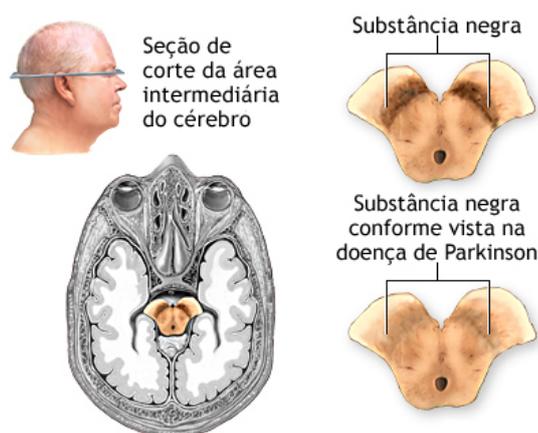
## 7 RESULTADOS EM CONTEXTO MULTICANAL: CLASSIFICAÇÃO DA DOENÇA DE PARKINSON

Neste capítulo, examinamos a aplicação do descritor de sinais em amostras de pacientes saudáveis e diagnosticados com a Doença de Parkinson, especialmente considerando a classificação multicanal com sinais de Eletroencefalograma (EEG).

### 7.1 A doença de Parkinson

A doença de Parkinson (DP) é uma condição neurodegenerativa crônica que afeta o sistema nervoso central, caracterizada principalmente por distúrbios motores. Ela resulta da degeneração progressiva de células nervosas na região do cérebro responsável pelo controle dos movimentos, conhecida como substância negra. Os sintomas distintivos incluem tremor involuntário, rigidez muscular, bradicinesia (lentidão nos movimentos), instabilidade postural e outros sintomas não motores, como distúrbios do sono e alterações cognitivas (ALJALAL *et al.*, 2022). A Figura 52 apresenta uma ilustração da doença de Parkinson na substância negra (POLAT; NOUR, 2020).

Figura 52 – Estado da doença de Parkinson ilustrando a degradação da substância negra, associada ao controle de movimentos. Adaptado de Polat e Nour (2020).



Fonte: Elaborada pelo autor.

Segundo Hamidi *et al.* (2023), embora o exame de tomografia computadorizada seja o mais utilizado com altas taxas de assertividade diagnóstica, estudos recentes destacam a crescente importância da análise de EEG, ressaltando suas taxas de acerto e sensibilidade no diagnóstico. Neste sentido, diversas abordagens envolvendo visão computacional, estudos estatísticos e aprendizado de máquina, tem ganhado destaque na pesquisa e desenvolvimento de

sistemas de apoio ao diagnóstico médico (SUGDEN; DIAMANDIS, 2023).

Neste sentido, Aljalal *et al.* (2022) ressalta que o exame de EEG é capaz de detectar padrões específicos de atividade elétrica cerebral associados à Doença de Parkinson. Alterações na sincronização e na amplitude das ondas cerebrais, especialmente nas faixas de frequência, já Hamidi *et al.* (2023) destaca que o EEG é sensível às oscilações corticais, que podem ser alteradas na Doença de Parkinson. Essas oscilações estão relacionadas a processos como controle motor e processamento cognitivo, que fornecem detalhes quanto ao comprometimento neural.

Dentro deste cenário, a metodologia proposta neste estudo apresenta combina o descritor de sinais baseado em grafos direcionados para a análise espectral dos sinais provenientes dos eletrodos, em exames de EEG. Visando melhorar a interpretabilidade dos sinais em conjunto com aprendizado de máquina para a classificação da doença. Diante dos resultados, esta abordagem se mostra promissora e visa contribuir com a superação das limitações na correlação direta entre os sintomas observados nos pacientes e os dados gerados pelo EEG, fomentando os avanços para o diagnóstico claro da doença.

## **7.2 Trabalhos Relacionados**

No primeiro capítulo deste trabalho, mais precisamente na Seção 1.3.2.2, exploramos distintas estratégias de extração de atributos em análises que envolvem múltiplos canais de sinais de EEG. Essas abordagens constituem uma contribuição essencial para a pesquisa realizada nesta seção. Além disso, apresentamos os estudos relacionados que servirão como base comparativa para os resultados obtidos nesta aplicação.

## **7.3 Amostras de pacientes**

Nas seções a seguir, serão apresentados dois conjuntos de dados, construídos sobre sinais de EEG abrangendo o cenário de extração de atributos para múltiplos canais. Em cada conjunto são apresentados os detalhes de número de amostras e eletrodos utilizados na aquisição dos sinais.

### **7.3.1 Dataset I**

Este estudo emprega um conjunto de dados obtido do Open Neuro (ROCKHILL *et al.*, 2020), composto por 16 indivíduos saudáveis (HC) e 15 pacientes diagnosticados com

doença de Parkinson (PD), não medicados. De acordo com Rockhill *et al.* (2020), este conjunto de dados é amplamente reconhecido e avaliado para estudos sobre o Parkinson, apresentando uma distribuição abrangente de características. O conjunto destaca-se por manter uma frequência de amostragem de 512 Hz, garantindo uma aquisição de dados em alta resolução. Além disso, os dados de cada paciente foi coletado por 32 canais (Fp1, AF3, F7, F3, FC1, FC5, T7, C3, CP1, CP5, P7, P3, Pz, PO3, O1, Oz, O2, PO4, P4, P8, CP6, CP2, C4, T8, FC6, FC2, F4, F8, AF4, Fp2, Fz, Cz). A escolha do cenário sem a influência de medicamentos, visa fornecer um comparativo sobre as características neurofisiológicas dos pacientes com Parkinson quando não estão sob tratamento medicamentoso.

### 7.3.2 Dataset II

A *University of Iowa* (UI) disponibilizou este conjunto de dados, formado por amostras de EEG para estudos quanto a classificação de pacientes com a doença de Parkinson (CAVANAGH *et al.*, 2018; SINGH *et al.*, 2020). O conjunto é composto pelos registros de EEG 14 pacientes, em repouso, diagnosticados com PD, e 14 indivíduos saudáveis (8 mulheres e 6 homens em ambos os grupos), apresentando uma média de idade de  $70,5 \pm 8,7$  anos. Os sinais de EEGs foram capturados na faixa de 0,1 a 100 Hz por meio de eletrodos de prata/cloreto de prata sinterizado, utilizando o sistema Brain Vision de 64 canais com uma frequência de amostragem de 500 Hz. Foram considerados neste trabalho sinais de de EEG contendo 10 canais, sendo eles: TP8, P6, FC5, PO8, O2, CP5, C2, C6, FT10, Cz. Embora a referência online tenha sido ajustada para o canal Pz como linha de base, é importante notar que o canal Pz está ausente neste conjunto de dados. O período médio de gravação por participante foi de  $3,14 \pm 1,06$  minutos.

## 7.4 Metodologia

A metodologia de extração de atributos dos sinais de EEG segue exatamente os mesmos passos para a construção do tensor espectral, considerando marcadores sobre as componentes espectrais, apresentados no capítulo 6, e tem como objetivo exatamente explorar a capacidade de generalização da metodologia em outro contexto de aplicação, além de avaliar o extrator considerando dois conjuntos de pacientes distintos.

### 7.4.1 Avaliação de Desempenho

Para avaliar o desempenho do método proposto quanto a tarefa de classificação das amostras, esta metodologia considera as seguintes métricas: Precisão (Acc), Precisão (Pre), Recall (Rec) e o F1-Score. Ademais, uma análise considerando o tempo de treinamento bem como de inferência em novos dados é levada em consideração.

## 7.5 Resultados e discussão

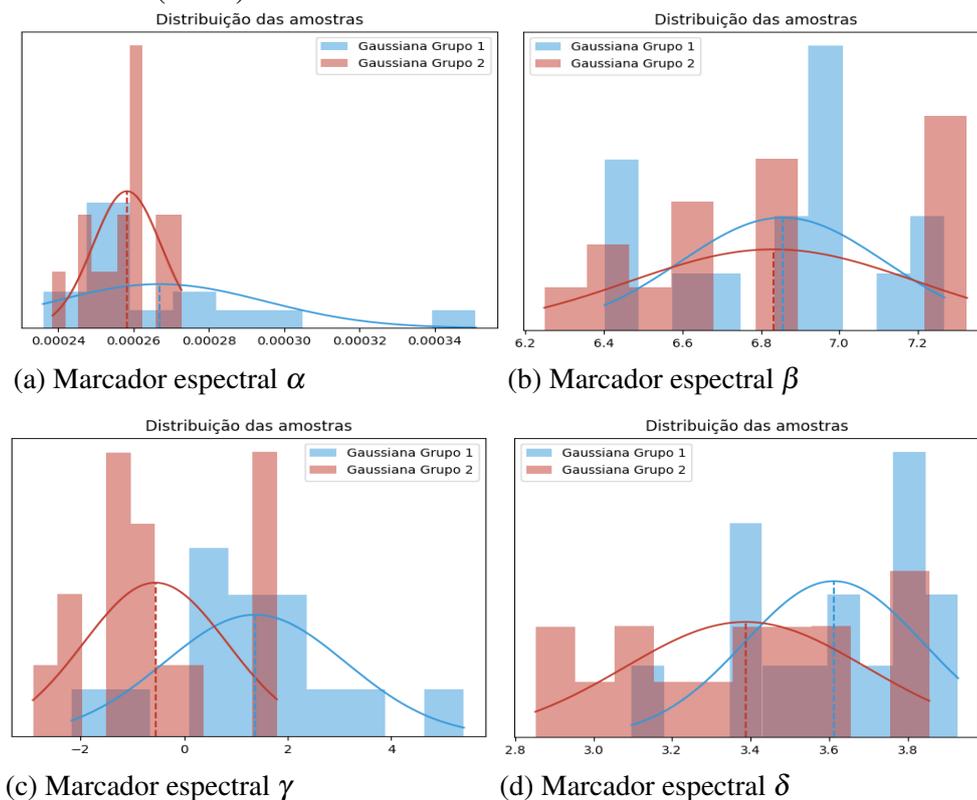
O método proposto utilizou diferentes arquiteturas de redes neurais tendo como entrada tensores tridimensionais. Essas arquiteturas são apresentadas no Capítulo 2. A etapa de classificação tem como base os dados extraídos pela metodologia proposta e usa o processo de validação cruzada *Stratified K-Folds*, com  $K = 5$ . Esse processo de validação cruzada é bastante utilizado em artigos mais recentes da literatura, principalmente em tarefas de classificação de doenças (GANSER *et al.*, 2021; GAMBOA *et al.*, 2022). Além disso, o ambiente de testes, bem como os modelos de classificação foram configurados de modo similar ao descrito na aplicação do Capítulo 6.

### 7.5.1 Teste estatístico ANOVA

A Figura 53 apresenta a distribuição das amostras após a extração, contemplando quatro diferentes situações de acordo com os marcadores espectrais. Os resultados apresentados referem-se a extração no *dataset I* e considera as classes HC vs. PD. Com base na distribuição apresentada no histograma, é possível observar que embora os marcadores  $\alpha$  e  $\beta$  tenham apresentado resultados insatisfatórios quanto a distância entre as médias dos grupos, os marcadores  $\gamma$  e  $\delta$  apresentam melhores resultados.

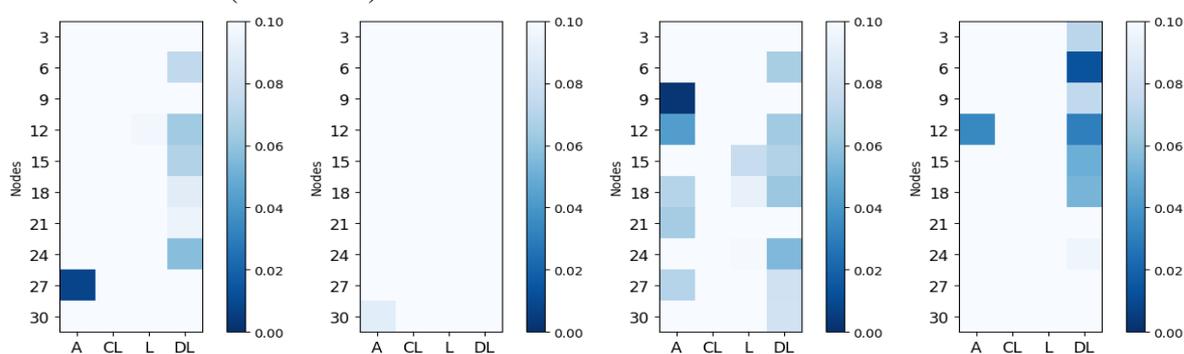
De modo similar, a Figura 54, exibe a distribuição das amostras extraídas e as curvas gaussianas em relação ao *dataset II*. Neste cenário, entretanto, apenas o marcador espectral *delta* sugere resultados melhores quanto a separação entre classes saudáveis e doentes. Por outro lado, a Figura 55 apresenta o valor- $p$  para as amostras extraídas das classes PD vs. HC, do *dataset I*. São apresentados cenários conforme o tipo de marcador. As linhas representam as diferentes configurações de número de vértices e as colunas, as matrizes A, CL, L e DL, que representam as diferentes perspectivas do grafo direcionado.

Figura 53 – Histograma das amostras extraídas para as classes PD vs. HC do *dataset* I. ( $L=12$ )



Fonte: Elaborada pelo autor.

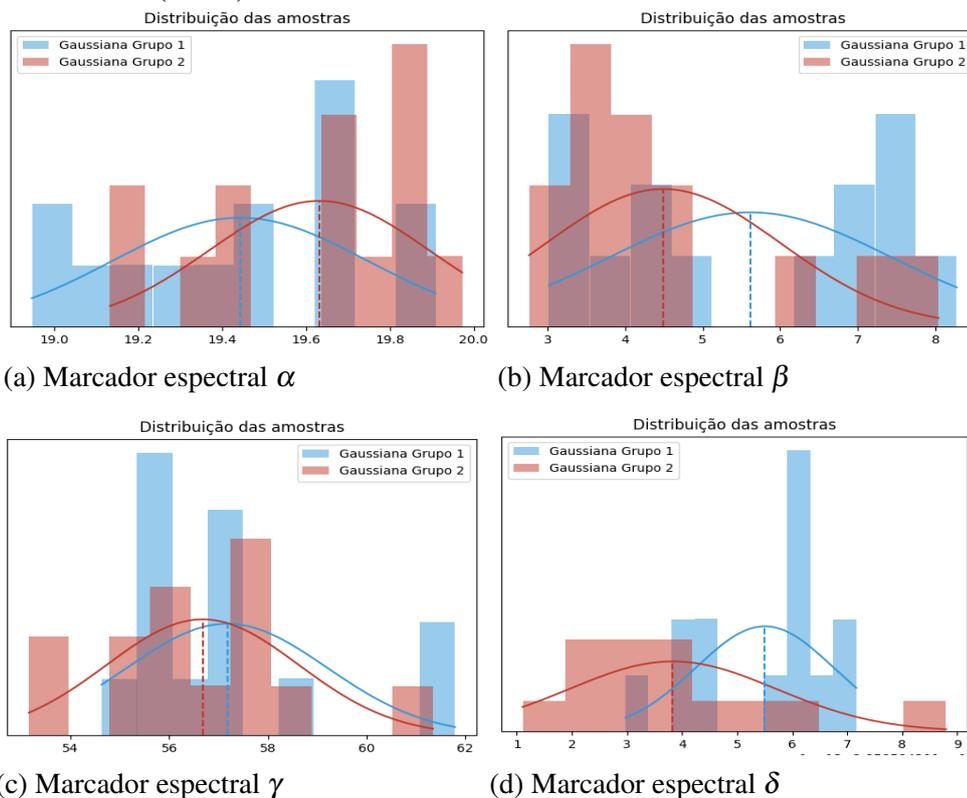
Figura 55 – Resultado do valor- $p$  no teste ANOVA para discriminação no *dataset* I (PD vs HC)



Fonte: Elaborada pelo autor.

É possível observar que as matrizes A e DL apresentam valor- $p$  inferiores ao nível de significância  $\alpha$  (0.05), como pode ser observado conforme escala ao lado. O que indica rejeição da hipótese nula e assim, indicam potencial de separação das classes. Já a matriz CL e L, entretanto, não apresentam valor- $p$  abaixo do nível de significância e, portanto, não rejeitamos a hipótese nula para estes casos. Destaca-se também que os marcadores  $\gamma$  e  $\delta$  indicam melhores

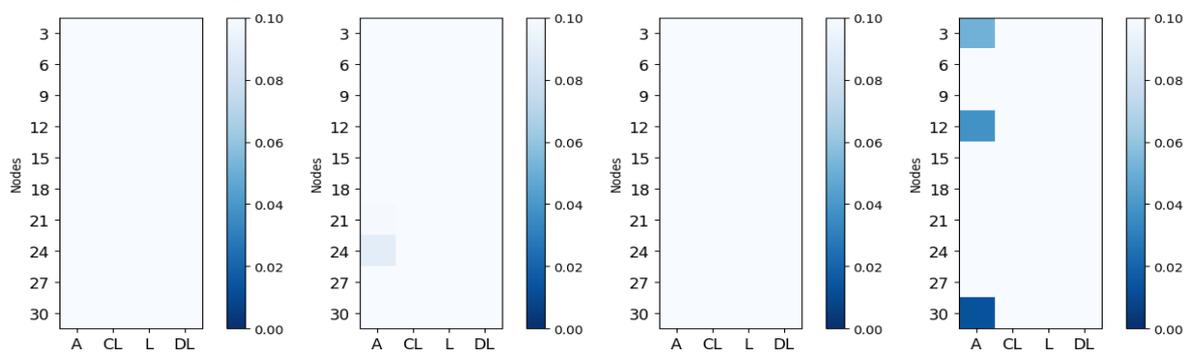
Figura 54 – Histograma das amostras extraídas para as classes PD vs. HC do *dataset* II. ( $L=12$ )



Fonte: Elaborada pelo autor.

resultados. Por outro lado, quanto ao *dataset* II, apenas a matriz de adjacência apresentou valores que indicam rejeição da hipótese nula, e indica deste modo, seu potencial para discriminação.

Figura 56 – Resultado do valor- $p$  no teste ANOVA para discriminação no *dataset* II (PD vs HC)



Fonte: Elaborada pelo autor.

Por outro lado, quanto aos resultados de classificação do tensor espectral, são apresentados a Tabela 16 apresenta o resultado da extração, considerando o modelo que alcançou

as melhores taxas de acerto. De acordo com a tabela, é possível notar que a combinação do modelo residual com a metodologia de extração proposta alcança índices de acerto promissores, alcançando valores Acurácia (99.9%), Precisão (99.9%) e AUC (99.9%) relevantes em relação as amostras do *dataset I*, correspondendo a classificação de amostras PD vs. HC.

Na comparação dos resultados obtidos pelo método proposto com outros métodos, a Tabela 16 revela que a metodologia proposta por Hamidi *et al.* (2023) apresenta o menor índice de precisão (ACC), atingindo 90.32%. Este desempenho inferior pode estar relacionado às dificuldades na generalização do modelo para diferentes distribuições estocásticas das amostras. Já os métodos propostos por Aljalal *et al.* (2022) e Loh *et al.* (2021), por outro lado, demonstram maior consistência nos resultados, alcançando índices de acurácia de 99.41% e 99.46%, respectivamente.

Vale destacar que o estudo de Aljalal *et al.* (2022) apresenta algumas limitações, como a falta de exploração da seleção de canais e uma análise mais aprofundada sobre a robustez da classificação. Essas limitações ressaltam áreas que demandam investigação adicional para aprimorar a aplicabilidade clínica dos métodos propostos. Além disso, o processo de classificação utiliza o k-vizinhos mais próximos (KNN), que pode não ser tão eficiente para grandes conjuntos.

Já a metodologia proposta por Loh *et al.* (2021) demonstra ser computacionalmente custosa, considerando a arquitetura da CNN proposta. Isso pode trazer limitações para aplicações embarcadas ou com requisito de tempo real. A metodologia proposta neste trabalho, entretanto, embora utilize CNN do tipo residual, apresentou um tempo de treino e teste bastante promissores, conforme Tabela 16.

Em contrapartida, a metodologia deste trabalho alcançou, em média, 99% em todas as métricas de avaliação, superando o trabalho de Hamidi *et al.* (2023) em 9%. Com resultados ligeiramente superiores também aos métodos de Aljalal *et al.* (2022) e Loh *et al.* (2021). Isso confirma o potencial da metodologia de extração baseada em grafo e que sua combinação com rede do tipo ResNet apresentou resultados significativos na análise multicanal do sinal de EEG.

Conforme os resultados do *dataset II*, Tabela 16), observa-se que o método proposto continua apresentando índices promissores para as métricas de Acurácia (96,70%), Precisão (97,22%), Recall (96,70%) e F1-Score (96,63%), superando a abordagem de Sugden e Diamandis (2023), no mesmo conjunto de dados, em até 10% de acurácia. Sendo também superior aos resultados de Delfan *et al.* (2023), em mais de 20% no índice de precisão. Por fim, quanto ao tempo de predição, a metodologia apresentou um custo de tempo pouco superior a 7 segundos,

enquanto a metodologia de Anjum *et al.* (2020) classifica a um tempo de 120 segundos.

Tabela 16 – Resultados para o melhor desempenho do método proposto nas amostras do dataset I e II, para classificação do Parkinson em relação a pacientes saudáveis.

Dataset	Subset	Descritor	Classificador	ACC	Pre.	Recall	F1-Score	Treino(s)	Teste(s)
Dataset I	PD vs HC	<b>Proposto (<math>\delta</math>-12 vértices)</b>	ResNet-152	<b>99.95</b>	<b>99.92</b>	<b>99.90</b>	<b>99.91</b>	<b>62.46</b>	<b>17.03</b>
		Hamidi <i>et al.</i> (2023)	CNN	90.32	-	-	-	-	-
		Aljalal <i>et al.</i> (2022)	KNN	99.41	-	-	99.40	-	-
		Loh <i>et al.</i> (2021)	CNN	99.46	99.48	-	99.46	-	-
Dataset II	PD vs HC	<b>Proposto (<math>\delta</math>-12 vértices)</b>	ResNet-50	<b>96.70</b>	<b>97.22</b>	<b>96.70</b>	<b>96.63</b>	<b>8.69</b>	<b>7.27</b>
		Sugden e Diamandis (2023)	CNN	82.80	-	-	84.11	-	-
		Delfan <i>et al.</i> (2023)	CNN, BiGRU	73.20	69.8	77.30	73.90	-	-
		Anjum <i>et al.</i> (2020)	SVM	85.70	85.70	-	-	-	120.00

Fonte: Elaborada pelo autor.

Finalmente, a Tabela 17 nos fornece uma visão geral dos métodos de classificação da doença de Parkinson. Isso oferece uma breve comparação dos cenários avaliados neste estudo, apresentando os resultados alcançados em trabalhos recentes no estado da arte em contraste com o que está sendo proposto.

## 7.6 Conclusões do Capítulo

Neste capítulo a metodologia para descrição de sinal foi aplicada no cenário de classificação da doença de Parkinson, a terceira doença neurodegenerativa sob a qual a metodologia foi avaliada. Dois conjuntos diferentes de pacientes foram avaliados com a abordagem proposta, configurada de modo semelhante àquela da aplicação com a doença de Alzheimer, Capítulo 6. Quanto aos índices de acerto, a metodologia apresentou resultados promissores de Acurácia (DB I: 99%, DB II 96.7%), Precisão (DB I: 99.92%, DB II 97.22%) e F1-Score (DB I: 99.91%, DB II 96.63%). Os resultados indicaram relevância para aderência do tensor espectral com o modelo de rede neural residual ResNet. Além de um tempo de predição inferior a 10 segundos. Esses resultados indicam que a abordagem de extração deste trabalho enfatiza de modo promissor as características mais relevantes do sinal de EEG, contribuindo com o processo de classificação.

Tabela 17 – Sumário comparativo com trabalhos relacionados a classificação de Parkinson.

Cenário	Estudo	Ano	Subset	Acurácia / AUC
PD vs HC	<b>Proposto</b>	<b>2023</b>	I. 15 PD, 16 HC II. 14 PD, 14 HC	<b>I. 99.95 / 99.00</b> <b>II. 96.70 / -</b>
	Hamidi <i>et al.</i> (2023)	2023	15 PD, 16 HC	90.32 / -
	Aljalal <i>et al.</i> (2022)	2022	15 PD, 16 HC	99.41 / -
	Loh <i>et al.</i> (2021)	2021	15 PD, 16 HC	99.46 / -
	Sugden e Diamandis (2023)	2023	14 PD, 14 HC	82.80 / 87.73
	Delfan <i>et al.</i> (2023)	2023	14 PD, 14 HC	73.20 / -
	Anjum <i>et al.</i> (2020)	2020	14 PD, 14 HC	85.70 / 85.20
	Shah <i>et al.</i> (2020)	2020	28 PD	99.22 / -

Fonte: Elaborada pelo autor.

## 8 CONCLUSÕES E TRABALHOS FUTUROS

Neste trabalho foi apresentada uma nova metodologia para representação do sinal de uma dimensão. Fundamentada com base na teoria dos grafos, a metodologia propõe a representação do sinal em um digrafo, cujos vértices representam níveis dos coeficientes de magnitude, já as arestas indicam as transições que são identificadas com base em quatro diferentes marcadores espectrais propostos, que processam pontos de máximos, pontos de mínimos, bem como pontos de inflexão. Tais marcadores contribuem para uma análise robusta a ruídos e de baixo custo computacional, já que os marcadores espectrais são definidos sobre operações matemáticas básicas.

O digrafo criado é representado sob três diferentes perspectivas: Matriz de Adjacência, Matriz Laplaciana e Matriz Laplaciana Combinatória Direcionada, consolida a análise em uma estrutura tensorial, chamada de tensor espectral. O tensor espectral encapsula as três perspectivas diferentes do grafo, que constituem cada uma das três matrizes. Esta proposta fomenta sua utilização com técnicas recentes de redes neurais baseadas em tensores, como redes neurais convolucionais, o que abre caminho para o uso de técnicas recentes como transferência de aprendizado.

A metodologia também mostrou-se flexível quanto ao número de sinais, sendo avaliada em aplicações monocanais, com sinais de voz, para extração de característica de pacientes com a Doença de Huntington, mas também em cenário multicanal, com sinais de eletroencefalograma, neste caso, avaliou-se o método em dois diferentes contextos: Doença de Alzheimer e Doença de Parkinson.

Os resultados indicaram que a metodologia proposta mostra potencial para representação de classes, tanto em contexto de problemas binários quanto em cenários multiclasse, como apresentado no Capítulo 6. O potencial de discriminação foi avaliado sob teste estatístico ANOVA, indicando cenários de rejeição da hipótese nula em todas as aplicações consideradas.

Esta tese também propôs a aplicação da metodologia de extração em três diferentes cenários com dados reais, de pacientes com diferentes doenças neurodegenerativas. A primeira aplicação propôs o processamento de sinais de voz, registrado por pacientes enquanto faziam a leitura de dois diferentes poemas. Os pacientes foram previamente classificados como saudáveis e portadores da Doença de Huntington. Os sinais de voz foram submetidos a extrações utilizando extratores estatísticos, baseados em Wavelets, em componentes espectrais a partir da série de Fourier e componentes Cepstrais; além destes, foram considerados também extratores

especializados em sinais de voz, por fim, o extrator proposto foi utilizado no mesmo conjunto. Os resultados indicaram potencial promissor da abordagem de extração de atributos do sinal de voz, indicando taxas de acurácia de 91,94%, com precisão de 92% e F1-Score de 91,31%. Tais indicadores sugerem que a metodologia proposta apresentou potencial competitivo para esta aplicação.

Por outro lado, com vistas a expandir as avaliações da abordagem proposta, avaliou-se também cenários multicanais. No primeiro cenário multicanal, este trabalho sugeriu aplicação do método com amostras de sinais provenientes do exame de eletroencefalograma (EEG), utilizando diferentes quantidades de canais, oriundos dos eletrodos, em um cenário de aplicação para classificação de pacientes portadores da Doença de Alzheimer. Neste cenário foram avaliados cinco diferentes *dataset*, sendo três aplicados no cenário de classificação binária (doentes vs saudáveis), e outros dois *dataset* para cenário multiclasse (Doença de Alzheimer vs. Saudáveis vs. Outras Patologias). Nos cenários de classificação binária, a metodologia apresentou índices de acurácia, precisão, F1-Score e AUC próximos a 100%, já no cenário de classificação multiclasse, próximos a 90%, ainda assim, superando abordagens que utilizaram o mesmo conjunto de dados nos experimentos.

Por fim, uma terceira aplicação proposta, contemplando outra doença neurodegenerativa, foi a classificação de pacientes saudáveis em relação a pacientes portadores da Doença de Parkinson. Neste cenário foram avaliados dois diferentes *datasets* já avaliados em estudos anteriores da literatura. Nesta aplicação, os índices de acurácia e AUC indicaram valores de 99,95% e 99%, respectivamente, para o *dataset* I. Já para o *dataset* II, a acurácia alcançou 96,70%. Em ambos os cenários a metodologia apresentou resultados aproximados ou superiores a outras abordagens que avaliaram as amostras dos mesmos *datasets*.

Diante dos resultados apresentados nos experimentos, quanto a tarefa de classificação, é possível concluir que os tensores espectrais propostos apresentaram resultados superiores a 90% em todas as aplicações consideradas. Neste contexto, observou-se uma boa aderência dos tensores espectrais quando combinados com a arquitetura ResNet, indicando resultados iguais ou superiores as demais redes também avaliadas. Este resultado pode indicar que a estratégia residual aplicada nas camadas da arquitetura destas redes, pode conduzir melhor o treinamento da combinação do tensor em suas diferentes perspectivas embutidas, contribuindo positivamente para o treinamento dos neurônios mesmo com gradientes próximos.

Por fim, a proposta mostrou-se competitiva em relação as tecnologias desenvolvidas

especificamente para classificação das doenças neurodegenerativas consideradas, indicando que a proposta do trabalho é promissora na análise dos sinais dos pacientes e pode contribuir para o desenvolvimento de tecnologias de auxílio ao diagnóstico.

### **8.0.1 Trabalhos Futuros**

Estudos futuros dessa metodologia visam envolver mais amostras, expandindo o número de pacientes afim de experimentar o potencial da proposta de extração em um conjunto de dados mais heterogêneo. Além disso, avaliar a metodologia de classificação utilizando outras abordagens de amostragem como a metodologia *Leave-One-Subject-Out* (LOSO) e *Leave-One-Out Cross-Validation* (LOOCV) que embora computacionalmente caras, podem fornecer mais detalhes quanto as métricas avaliadas. Sob o ponto de vista de pesquisas aplicadas, elencamos algumas linhas para continuidade dos estudos, a saber:

- Aplicar operações de grafo, tais como identificar árvore geradora mínima e produzir um tensor espectral com base nesta nova estrutura, avaliando os impactos quanto a representatividade;
- Aplicar operações de poda no grafo, uma vez que as frequências ruidosas tendem a se concentrar em um nível de magnitude, portanto, em um vértice do grafo. Com vistas a reduzir ainda mais a influência do ruído do sinal durante a fase de classificação do tensor;
- Avaliar a eficácia desta proposta em outras doenças neurodegenerativas visando identificar pacientes nos estados iniciais da doença, tais como Esclerose Lateral Amiotrófica (ELA), Atrofia Espinal Muscular, afim de contribuir com o diagnóstico precoce e tratamento mais efetivo;
- Avaliar o impacto de tensores espectrais quanto a compressão dos dados originalmente representados em sinais de voz e em sinais de EGG, sob o ponto de vista da redução de espaço de armazenamento ocupado, visando uma proposta de uso para compressão e armazenagem de dados;
- Avaliar o potencial de extração da metodologia proposta em aplicações de tempo real com mudanças constantes, como ocorre na análise de gráficos financeiros, onde o valor da ação ou da moeda altera constantemente e apresenta um ambiente de sinal bastante heterogêneo;
- Expandir o algoritmo do processamento de sinais de uma dimensão para processamento de sinais de duas dimensões, bem como imagens, visando avaliar sua aplicação como extrator de características bidimensional.

## REFERÊNCIAS

- ALBAWI, S.; MOHAMMED, T. A.; AL-ZAWI, S. Understanding of a convolutional neural network. **2017 international conference on engineering and technology (ICET)**, IEEE, Antalya, Turquia, p. 1–6, 2017.
- ALJALAL, M.; ALDOSARI, S. A.; ALSHARABI, K.; ABDURRAQEEB, A. M.; ALTURKI, F. A. Parkinson's disease detection from resting-state eeg signals using common spatial pattern, entropy, and machine learning techniques. **Diagnostics**, MDPI, Basel, Switzerland, v. 12, n. 9, p. 1–24, 2022.
- ALMEIDA, J. S.; FILHO, P. P. R.; CARNEIRO, T.; WEI, W.; DAMAŠEVIČIUS, R.; MASKELIŪNAS, R.; ALBUQUERQUE, V. H. C. de. Detecting parkinson's disease with sustained phonation and speech signals using machine learning techniques. **Pattern Recognition Letters**, Elsevier, v. 125, p. 55–62, 2019.
- ALTAHERI, H.; MUHAMMAD, G.; ALSULAIMAN, M.; AMIN, S. U.; ALTUWAIJRI, G. A.; ABDUL, W.; BENCHERIF, M. A.; FAISAL, M. Deep learning techniques for classification of electroencephalogram (eeg) motor imagery (mi) signals: A review. **Neural Computing and Applications**, Springer, v. 35, n. 20, p. 14681–14722, 2023.
- ANJUM, M. F.; DASGUPTA, S.; MUDUMBAL, R.; SINGH, A.; CAVANAGH, J. F.; NARAYANAN, N. S. Linear predictive coding distinguishes spectral eeg features of parkinson's disease. **Parkinsonism & related disorders**, Elsevier, v. 79, p. 79–85, 2020.
- APPAKAYA, S. B.; PRATIHAR, R.; SANKAR, R. Parkinson's disease classification framework using vocal dynamics in connected speech. **Algorithms**, MDPI, Basel, Switzerland, v. 16, n. 11, p. 1–19, 2023.
- ARAÚJO, T.; TEIXEIRA, J. P.; RODRIGUES, P. M. Smart-data-driven system for alzheimer disease detection through electroencephalographic signals. **Bioengineering**, MDPI, v. 9, n. 4, p. 141, 2022.
- BANG-JENSEN, J.; GUTIN, G. Z. Digraphs: theory, algorithms and applications. **Monographs in Mathematics**, Springer, Nova York, v. 2, 2008.
- BATTAGLIA, S.; SCHMIDT, A.; HASSEL, S.; TANAKA, M. Case reports in neuroimaging and stimulation. **Frontiers in Psychiatry**, Frontiers Media SA, v. 14, 2023.
- BEKELE, A. Cooley-tukey fft algorithms. **Advanced algorithms**, Ottawa, v. 19, n. 90, p. 297–301, 2016.
- BOVOLENTA, T. M.; SILVA, S. M. C. de A.; SABA, R. A.; BORGES, V.; FERRAZ, H. B.; FELICIO, A. C. Average annual cost of parkinson's disease in são paulo, brazil, with a focus on disease-related motor symptoms. **Clinical interventions in aging**, Taylor & Francis, p. 2095–2108, 2017.
- BREIMAN, L. Random forests. **Machine learning**, Springer, v. 45, p. 5–32, 2001.
- BURRUS, C. S.; PARKS, T. Convolution algorithms. **Citeseer: New York**, Citeseer, v. 6, p. 15, 1985.

CAVANAGH, J. F.; KUMAR, P.; MUELLER, A. A.; RICHARDSON, S. P.; MUEEN, A. Diminished eeg habituation to novel events effectively classifies parkinson's patients. **Clinical Neurophysiology**, v. 129, n. 2, p. 409–418, 2018.

CERVATO, M. **Doença de Huntington**. São Paulo: [s.n.], 2023. Disponível em: <https://blog.varsomics.com/doenca-de-huntington>. Acesso em: 15 nov. 2023.

CHEN, Y.; WANG, H.; ZHANG, D.; ZHANG, L.; TAO, L. Multi-feature fusion learning for alzheimer's disease prediction using eeg signals in resting state. **Frontiers in neuroscience**, Frontiers Media SA, v. 17, 2023.

CHUNG, F. Laplacians and the cheeger inequality for directed graphs. **Annals of Combinatorics**, Springer, v. 9, p. 1–19, 2005.

COOLEY, J. W.; TUKEY, J. W. An algorithm for the machine calculation of complex fourier series. **Mathematics of computation**, v. 19, n. 90, p. 297–301, 1965.

COVER, T.; HART, P. Nearest neighbor pattern classification. **IEEE transactions on information theory**, IEEE, v. 13, n. 1, p. 21–27, 1967.

CROVATO, C. D. P.; SCHUCK, A. The use of wavelet packet transform and artificial neural networks in analysis and classification of dysphonic voices. **IEEE Transactions on Biomedical Engineering**, IEEE, v. 54, n. 10, p. 1898–1900, 2007.

DELFAN, N.; SHAHSAVARI, M.; HUSSAIN, S.; DAMAŠEVIČIUS, R.; ACHARYA, U. R. A hybrid deep spatio-temporal attention-based model for parkinson's disease diagnosis using resting state eeg signals. **arXiv preprint**, Cornell University, Nova York, v. 4, 2023.

DIESTEL, R.; DIESTEL, R. Extremal graph theory. **Graph theory**, Springer, p. 173–207, 2017.

DIESTEL, R.; KUHN, D. Graph-theoretical versus topological ends of graphs. **Journal of Combinatorial Theory, Series B**, Elsevier, v. 87, n. 1, p. 197–206, 2003.

DOGAN, S.; BAYGIN, M.; TASCI, B.; LOH, H. W.; BARUA, P. D.; TUNCER, T.; TAN, R.-S.; ACHARYA, U. R. Primate brain pattern-based automated alzheimer's disease detection model using eeg signals. **Cognitive Neurodynamics**, Springer, v. 17, n. 3, p. 647–659, 2023.

DONG, X.; THANOU, D.; FROSSARD, P.; VANDERGHEYNST, P. Learning laplacian matrix in smooth graph signal representations. **IEEE Transactions on Signal Processing**, IEEE, v. 64, n. 23, p. 6160–6173, 2016.

DORSEY, E.; SHERER, T.; OKUN, M. S.; BLOEM, B. R. The emerging evidence of the parkinson pandemic. **Journal of Parkinson's disease**, IOS Press, v. 8, n. s1, p. 3–8, 2018.

DUDA, R. O.; HART, P. E. *et al.* Pattern classification. **John Wiley Sons**, 2006.

EULER, L. Solutio problematis ad geometriam situs pertinentis. **Commentarii academiae scientiarum Petropolitanae**, São Petersburgo, p. 128–140, 1741.

EYBEN, F.; WÖLLMER, M.; SCHULLER, B. OpenSMILE: The Munich versatile and fast open-source audio feature extractor. **18th ACM international conference on Multimedia**, p. 1459–1462, 2010.

FISCON, G.; WEITSCHKE, E.; CIALINI, A.; FELICI, G.; BERTOLAZZI, P.; SALVO, S. D.; BRAMANTI, A.; BRAMANTI, P.; COLA, M. C. D. Combining eeg signal processing with supervised methods for alzheimer's patients classification. **BMC medical informatics and decision making**, BioMed Central, v. 18, n. 1, p. 1–10, 2018.

FONG, S. *et al.* Using hierarchical time series clustering algorithm and wavelet classifier for biometric voice classification. **BioMed Research International**, Hindawi, v. 2012, 2012.

FRIEDMAN, J. H. Greedy function approximation: a gradient boosting machine. **Annals of statistics**, JSTOR, p. 1189–1232, 2001.

GALLEZOT, C.; RIAD, R.; TITEUX, H.; LEMOINE, L.; MONTILLOT, J.; SLIWINSKI, A.; BAGNOU, J. H.; CAO, X. N.; YOUSSEF, K.; DUPOUX, E. *et al.* Emotion expression through spoken language in huntington disease. **Cortex**, Elsevier, v. 155, p. 150–161, 2022.

GAMBOA, P.; VARANDAS, R.; RODRIGUES, J.; CEPEDA, C.; QUARESMA, C.; GAMBOA, H. Attention classification based on biosignals during standard cognitive tasks for occupational domains. **Computers**, MDPI, v. 11, n. 4, p. 49, 2022.

GANSER, A.; HOLLAUS, B.; STABINGER, S. Classification of tennis shots with a neural network approach. **Sensors**, MDPI, v. 21, n. 17, p. 5703, 2021.

GARCÍA, M. J. V.; COBETA, I.; MARTÍN, G.; ALONSO-NAVARRO, H.; JIMENEZ-JIMENEZ, F. J. Acoustic analysis of voice in huntington's disease patients. **Journal of Voice**, Elsevier, v. 25, n. 2, p. 208–217, 2011.

GAUTHIER, S.; LEUZY, A.; RACINE, E.; ROSA-NETO, P. Diagnosis and management of alzheimer's disease: past, present and future ethical issues. **Progress in neurobiology**, Elsevier, v. 110, p. 102–113, 2013.

GEURTS, P.; ERNST, D.; WEHENKEL, L. Extremely randomized trees. **Machine learning**, Springer, v. 63, p. 3–42, 2006.

GUIMARÃES, M. T.; MEDEIROS, A. G.; ALMEIDA, J. S.; MARTIN, M. F. y; DAMAŠEVIČIUS, R.; MASKELIŪNAS, R.; MATTOS, C. L. C.; FILHO, P. P. R. An optimized approach to huntington's disease detecting via audio signals processing with dimensionality reduction. **2020 International Joint Conference on Neural Networks**, IEEE, p. 1–8, 2020.

HAMIDI, A.; YOUSEFI, M. *et al.* Classification of eeg signals to detect parkinsons disease using a computationally method. **arXiv preprint**, Cornell University, Nova York, v. 4, 2023.

HAN, J.; PEI, J.; TONG, H. Data mining: concepts and techniques. **Morgan Kaufmann Publishers**, 2022.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. H.; FRIEDMAN, J. H. The elements of statistical learning: data mining, inference, and prediction. **Springer Series in Statistics**, Springer, v. 2, 2009.

HE, K.; ZHANG, X.; REN, S.; SUN, J. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. **Proceedings of the IEEE international conference on computer vision**, IEEE, p. 1026–1034, 2015.

- HE, K.; ZHANG, X.; REN, S.; SUN, J. Deep residual learning for image recognition. **Proceedings of the IEEE conference on computer vision and pattern recognition**, IEEE, p. 770–778, 2016.
- HOSSEINI, M.-P.; HOSSEINI, A.; AHI, K. A review on machine learning for eeg signal processing in bioengineering. **IEEE reviews in biomedical engineering**, IEEE, v. 14, p. 204–218, 2020.
- HUANG, J. **Demência**. São Paulo: Manual MSD, 2023. Disponível em: <https://t.ly/1ErF5>. Acesso em: 13 jan. 2024.
- IOFFE, S.; SZEGEDY, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. **arXiv preprint**, Cornell University, Nova York, v. 4, 2015.
- IRAM, S.; FERGUS, P.; JUMEILY, D. A.; HUSSAIN, A.; RANGLES, M. A classifier fusion strategy to improve the early detection of neurodegenerative diseases. **International Journal of Artificial Intelligence and Soft Computing**, Inderscience Publishers, v. 5, n. 1, p. 23, 2015.
- JESUS, R. C. A. **Meio ambiente e doenças neurodegenerativas: traçando conexões com base em população do sudoeste sergipano, Brasil**. Tese (Doutorado) — Universidade Federal de Sergipe, 2022.
- JOHNSON, K. J.; SYNOVEC, R. E. Pattern recognition of jet fuels: comprehensive  $gc \times gc$  with anova-based feature selection and principal component analysis. **Chemometrics and Intelligent Laboratory Systems**, Elsevier, v. 60, n. 1-2, p. 225–237, 2002.
- KE, G.; MENG, Q.; FINLEY, T.; WANG, T.; CHEN, W.; MA, W.; YE, Q.; LIU, T.-Y. Lightgbm: A highly efficient gradient boosting decision tree. **Advances in neural information processing systems**, Neural Information Processing Systems, v. 30, 2017.
- KHALIFA, F.; SHALABY, A.; SOLIMAN, A.; ELASKARY, S.; REFAEY, A.; ABDELAZIM, M. Artificial intelligence-based computer-aided diagnosis applications for brain disorders from medical imaging data. **Frontiers in Neuroscience**, Frontiers, v. 17, p. 93–101, 2023.
- LEITHOLD, L. **O Cálculo Com Geometria Analítica**. São Paulo: Harbra, 1994. v. 2.
- LIU, S.; XIONG, Y.; CHEN, W.; CAO, H. Feature extraction of high frequency vocal fold paralysis voice based on stationary wavelets as data filtering for machine learning. **2nd International Conference on Artificial Intelligence and Computer Information Technology (AICIT)**, p. 1–6, 2023.
- LOGAN, B. Mel frequency cepstral coefficients for music modeling. **International Symposium on Music Information Retrieval, MUSIC IR**, Plymouth, Massachusetts, v. 270, n. 1, p. 11, 2000.
- LOH, H. W.; OOI, C. P.; PALMER, E.; BARUA, P. D.; DOGAN, S.; TUNCER, T.; BAYGIN, M.; ACHARYA, U. R. Gaborpdnet: Gabor transformation and deep neural network for parkinson's disease detection using eeg signals. **Electronics**, MDPI, v. 10, n. 14, p. 1740, 2021.
- LOPES, M.; CASSANI, R.; FALK, T. H. *et al.* Using cnn saliency maps and eeg modulation spectra for improved and more interpretable machine learning-based alzheimer's disease diagnosis. **Computational Intelligence and Neuroscience**, Hindawi, v. 2023, 2023.

LU, B.; SUN, J.; SUN, Y.; SHEN, A.; LI, Y.; GAO, X. A cross-device ensemble learning model for continuous cognition assessment based on resting-state eeg. **NPJ Aging**, Nature, 2023.

MA, N.; ZHANG, X.; ZHENG, H.-T.; SUN, J. Shufflenet v2: Practical guidelines for efficient cnn architecture design. **Proceedings of the European conference on computer vision (ECCV)**, Springer, p. 116–131, 2018.

MATHUR, P.; CHAKKA, V. K. Graph signal processing of eeg signals for detection of epilepsy. **7th International Conference on Signal Processing and Integrated Networks (SPIN)**, IEEE, Noida, India, p. 839–843, 2020.

MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **The bulletin of mathematical biophysics**, Springer, v. 5, p. 115–133, 1943.

MEGHDADI, A. H.; KARIĆ, M. S.; BERKA, C. EEG analytics: benefits and challenges of data driven EEG biomarkers for neurodegenerative diseases. **IEEE International Conference on Systems, Man and Cybernetics (SMC)**, IEEE, Bari, Italy, p. 1280–1285, 2019.

MELLONE, S.; MANCINI, M.; KING, L. A.; HORAK, F. B.; CHIARI, L. The quality of turning in parkinson's disease: a compensatory strategy to prevent postural instability? **Journal of neuroengineering and rehabilitation**, Springer, v. 13, p. 1–9, 2016.

MILTADOUS, A.; GIONANIDIS, E.; TZIMOURTA, K. D.; GIANNAKEAS, N.; TZALLAS, A. T. Dice-net: a novel convolution-transformer architecture for alzheimer detection in eeg signals. **IEEE Access**, IEEE, Noida, India, 2023.

MILTADOUS, A.; TZIMOURTA, K. D.; AFRANTOU, T.; IOANNIDIS, P.; GRIGORIADIS, N.; TSALIKAKIS, D. G.; ANGELIDIS, P.; TSIPOURAS, M. G.; GLAVAS, E.; GIANNAKEAS, N.; TZALLAS, A. T. A dataset of eeg recordings from: Alzheimer's disease, frontotemporal dementia and healthy subjects. **OpenNeuro**, 2023.

MISHRA, P.; SINGH, U.; PANDEY, C. M.; MISHRA, P.; PANDEY, G. Application of student's t-test, analysis of variance, and covariance. **Annals of cardiac anaesthesia**, Wolters Kluwer–Medknow Publications, v. 22, n. 4, p. 407, 2019.

MOLINA, G. N. G. **Direct brain-computer communication through scalp recorded EEG signals**. Tese (Doutorado) — École polytechnique fédérale de Lausanne, Lausanne, Suíça, 2004.

MOOTOO, X. S.; FOURS, A.; DINESH, C.; ASHKANI, M.; KISS, A.; FALTYN, M. Detecting alzheimer's disease in eeg data with machine learning and the graph discrete fourier transform. **medRxiv**, Cold Spring Harbor Laboratory Press, p. 23–31, 2023.

MURO-DE-LA-HERRAN, A.; GARCIA-ZAPIRAIN, B.; MENDEZ-ZORRILLA, A. Gait analysis methods: An overview of wearable and non-wearable systems, highlighting clinical applications. **Sensors**, Molecular Diversity Preservation International (MDPI), v. 14, n. 2, p. 3362–3394, 2014.

NAJEEB, S. M. M.; RIKABI, H. T. S. A.; ALI, S. M. Finding the discriminative frequencies of motor electroencephalography signal using genetic algorithm. **TELKOMNIKA (Telecommunication Computing Electronics and Control)**, v. 19, n. 1, p. 285–292, 2021.

NANDI, A. **Blind estimation using higher-order statistics**. Sail, Estados Unidos: Springer Science & Business Media, 1999.

NASIRI, H.; ALAVI, S. A. *et al.* A novel framework based on deep learning and anova feature selection method for diagnosis of covid-19 cases from chest x-ray images. **Computational intelligence and neuroscience**, Hindawi, v. 2022, 2022.

OPPENHEIM, A. V. **Discrete-time signal processing**. Noida, India: Pearson Education India, 1999.

PARIHAR, A.; SWAMI, P. D. Eeg classification of alzheimer's disease, frontotemporal dementia and control normal subjects using supervised machine learning algorithms on various eeg frequency bands. **International Journal Of Science Technology Management**, 2023.

PARRA, C. R.; TORRES, A. P.; SOTOS, J. M.; BORJA, A. L. Classification of moderate and advanced alzheimer's patients using radial basis function based neural networks initialized with fuzzy logic. **IRBM**, Elsevier, v. 44, n. 5, p. 98–105, 2023.

PEREZ, M.; JIN, W.; LE, D.; CARLOZZI, N.; DAYALU, P.; ROBERTS, A.; PROVOST, E. M. Classification of huntington disease using acoustic and lexical features. **Interspeech**, v. 2018, p. 1898, 2018.

PINEDA, A. M.; RAMOS, F. M.; BETTING, L. E.; CAMPANHARO, A. S. Quantile graphs for eeg-based diagnosis of alzheimer's disease. **Plos one**, Public Library of Science, Craiova, Romênia, v. 15, n. 6, 2020.

POLAT, K.; NOUR, M. Parkinson disease classification using one against all based data sampling with the acoustic features from the speech signals. **Medical Hypotheses**, Elsevier, v. 140, 2020.

PORSTEINSSON, A.; ISAACSON, R.; KNOX, S.; SABBAGH, M.; RUBINO, I. Diagnosis of early alzheimer's disease: clinical practice in 2021. **The journal of prevention of Alzheimer's disease**, Springer, v. 8, p. 371–386, 2021.

PRESS, W. H. **Numerical recipes: The art of scientific computing**. 3. ed. Cambridge, Inglaterra: Cambridge University Press, 2007.

PURI, D. V.; GAWANDE, J. P.; RAJPUT, J. L.; NALBALWAR, S. L. A novel optimal wavelet filter banks for automated diagnosis of alzheimer's disease and mild cognitive impairment using electroencephalogram signals. **Decision Analytics Journal**, Elsevier, 2023.

RAUGEI, S.; GERVASIO, F. L.; CARLONI, P. Dft modeling of biological systems. **Physica status solidi**, Wiley Online Library, Weinheim, Alemanha, v. 243, n. 11, p. 2500–2515, 2006.

RHODES, W. T. Acousto-optic signal processing: convolution and correlation. **Proceedings of the IEEE**, IEEE, v. 69, n. 1, p. 65–79, 1981.

RIAD, R.; TITEUX, H.; LEMOINE, L.; MONTILLOT, J.; BAGNOU, J. H.; CAO, X. N.; DUPOUX, E.; BACHOUUD-LÉVI, A.-C. Vocal markers from sustained phonation in huntington's disease. **arXiv preprint**, Cornell University, Nova York, v. 4, 2020.

RIOUL, O.; VETTERLI, M. Wavelets and signal processing. **IEEE signal processing magazine**, IEEE, v. 8, n. 4, p. 14–38, 1991.

ROCKHILL, A. P.; JACKSON, N.; GEORGE, J.; ARON, A.; SWANN, N. C. *et al.* Uc san diego resting state eeg data from patients with parkinson's disease. **OpenNeuro**, 2020.

ROMANA, A.; BANDON, J.; CARLOZZI, N.; ROBERTS, A.; PROVOST, E. M. Classification of manifest huntington disease using vowel distortion measures. **Interspeech**, NIH Public Access, v. 2020, p. 4966, 2020.

ROSENBLATT, F. The perceptron: a probabilistic model for information storage and organization in the brain. **Psychological review**, American Psychological Association, v. 65, n. 6, p. 386, 1958.

RUIZ-GÓMEZ, S. J.; GÓMEZ, C.; POZA, J.; GUTIÉRREZ-TOBAL, G. C.; TOLA-ARRIBAS, M. A.; CANO, M.; HORNERO, R. Automated multiclass classification of spontaneous eeg activity in alzheimer's disease and mild cognitive impairment. **Entropy**, MDPI, v. 20, n. 1, p. 35, 2018.

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. **Nature**, v. 323, n. 6088, p. 533–536, 1986.

SANDLER, M.; HOWARD, A.; ZHU, M.; ZHMOGINOV, A.; CHEN, L.-C. Mobilenetv2: Inverted residuals and linear bottlenecks. **Proceedings of the IEEE conference on computer vision and pattern recognition**, IEEE, p. 4510–4520, 2018.

SCHULLER, B.; STEIDL, S.; BATLINER, A.; BURKHARDT, F.; DEVILLERS, L.; MÜLLER, C.; NARAYANAN, S. The interspeech 2010 paralinguistic challenge. **INTERSPEECH 2010**, Sail, Estados Unidos, p. 2794–2797.

SCHULLER, B.; STEIDL, S.; BATLINER, A.; VINCIARELLI, A.; SCHERER, K.; RINGEVAL, F.; CHETOUANI, M.; WENINGER, F.; EYBEN, F.; MARCHI, E. *et al.* The interspeech 2013 computational paralinguistics challenge: Social signals, conflict, emotion, autism. **INTERSPEECH 2013**, International Speech Communication Association, 2013.

SCHULLER, B.; VALSTAR, M.; EYBEN, F.; MCKEOWN, G.; COWIE, R.; PANTIC, M. 2011 - the first international audio/visual emotion challenge. **Affective Computing and Intelligent Interaction: Fourth International Conference**, Springer, p. 415–424, 2011.

SCHÜTZE, H.; MANNING, C. D.; RAGHAVAN, P. **Introduction to information retrieval**. [S.l.]: Cambridge University Press Cambridge, 2008.

SHAH, S. A. A.; ZHANG, L.; BAIS, A. Dynamical system based compact deep hybrid network for classification of parkinson disease related eeg signals. **Neural Networks**, Elsevier, v. 130, p. 75–84, 2020.

SHAKARCHI, J. A. **The use of infrared thermal imaging as a marker of tissue perfusion and predictor of arteriovenous fistula outcomes**. Tese (Doutorado) — University of Birmingham, Birmingham, Reino Unido, 2017.

SIBILANO, E.; LASSI, M.; MAZZONI, A.; BEVILACQUA, V.; BRUNETTI, A. A deep learning framework for the classification of pre-prodromal and prodromal alzheimer's disease using resting-state eeg signals. **Applications of Artificial Intelligence and Neural Systems to Data Science**, Springer, p. 93–101, 2023.

SIMON, S.; HAYKIN, S. Neural networks and learning machines. **Pearson Upper Saddle River**, v. 3, 2009.

SINGH, A.; COLE, R. C.; ESPINOZA, A. I.; BROWN, D.; CAVANAGH, J. F.; NARAYANAN, N. S. Frontal theta and beta oscillations during lower-limb movement in parkinson's disease. **Clinical Neurophysiology**, v. 131, n. 3, p. 694–702, 2020.

SIJULY, S.; LI, Y.; ZHANG, Y. Eeg signal analysis and classification. **IEEE Trans Neural Syst Rehabil Eng**, Springer, v. 11, p. 141–144, 2016.

SOL, J.; AAEN, M.; SADOLIN, C.; BOSCH, L. T. Towards automated vocal mode classification in healthy singing voice—an xgboost decision tree-based machine learning classifier. **Journal of Voice**, Elsevier, 2023.

STAM, C. J.; JONES, B.; NOLTE, G.; BREAKSPEAR, M.; SCHELTENS, P. Small-world networks and functional connectivity in alzheimer's disease. **Cerebral cortex**, Oxford University Press, v. 17, n. 1, p. 92–99, 2007.

STHLE, L.; WOLD, S. *et al.* Analysis of variance (anova). **Chemometrics and intelligent laboratory systems**, Elsevier, v. 6, n. 4, p. 259–272, 1989.

SUBRAMANIAN, V.; KWON, N.; BRUECKNER, R.; BLAYLOCK, N.; O'CONNELL, H.; SIERRA, L.; ULLMAN, C.; HILDEBRAND, K.; LAGANIERE, S. Detecting manifest huntington's disease using vocal data. **Narrative**, v. 31, n. 11.3, p. 1–8, 2023.

SUGDEN, R. J.; DIAMANDIS, P. Generalizable electroencephalographic classification of parkinson's disease using deep learning. **Informatics in Medicine Unlocked**, Toronto, Canada, v. 42, 2023.

TAN, M.; CHEN, B.; PANG, R.; VASUDEVAN, V.; LE, Q. V. Mnasnet: Platform-aware neural architecture search for mobile. **CoRR**, abs/1807.11626, 2018.

TATUM, W. O. **Handbook of EEG interpretation**. [S.l.]: Springer Publishing Company, 2021.

THIES, T.; MÜCKE, D.; LOWIT, A.; KALBE, E.; STEFFEN, J.; BARBE, M. T. Prominence marking in parkinsonian speech and its correlation with motor performance and cognitive abilities. **Neuropsychologia**, Elsevier, v. 137, p. 300–306, 2020.

VAICIUKYNAS, E.; VERIKAS, A.; GELZINIS, A.; BACAUSKIENE, M. Detecting Parkinson's disease from sustained phonation and speech signals. **PLOS ONE**, Public Library of Science, v. 12, n. 10, p. 1–16, 10 2017.

VAPNIK, V. N. An overview of statistical learning theory. **IEEE transactions on neural networks**, IEEE, v. 10, n. 5, p. 988–999, 1999.

VÁSQUEZ-CORREA, J. C.; RIOS-URREGO, C. D.; ARIAS-VERGARA, T.; SCHUSTER, M.; RUSZ, J.; NÖTH, E.; OROZCO-ARROYAVE, J. R. Transfer learning helps to improve the accuracy to classify patients with different speech disorders in different languages. **Pattern Recognition Letters**, Elsevier, v. 150, p. 272–279, 2021.

VERAS, R. P.; CALDAS, C. P.; DANTAS, S. B.; SANCHO, L. G.; SICSU, B.; MOTTA, L. B. Demented elderly people living at home in rio de janeiro, brazil: Evaluation of expenditure on care. **Psychogeriatrics**, Wiley Online Library, v. 8, n. 2, p. 88–95, 2008.

VICCHIETTI, M. L.; RAMOS, F. M.; BETTING, L. E.; CAMPANHARO, A. S. Computational methods of eeg signals analysis for alzheimer's disease classification. **Scientific Reports**, Nature Publishing Group UK London, v. 13, n. 1, p. 8184, 2023.

WEST, D. B. *et al.* **Introduction to graph theory**. [S.l.]: Prentice hall Upper Saddle River, 2001. v. 2.

WIMO, A.; GUERCHET, M.; ALI, G.-C.; WU, Y.-T.; PRINA, A. M.; WINBLAD, B.; JÖNSSON, L.; LIU, Z.; PRINCE, M. The worldwide costs of dementia 2015 and comparisons with 2010. **Alzheimer's & Dementia**, Elsevier, v. 13, n. 1, p. 1–7, 2017.

WORLD HEALTH ORGANIZATION. **Global Dementia Observatory**. [S.l.]: World Health Organization, 2022. Disponível em: <https://t.ly/S4spc>. Acesso em: 13 jan. 2024.

XIAO, R.; HU, Q.; LI, J. Leak detection of gas pipelines using acoustic signals based on wavelet transform and support vector machine. **Measurement**, Elsevier, v. 146, p. 479–489, 2019.

XIE, S.; GIRSHICK, R.; DOLLÁR, P.; TU, Z.; HE, K. Aggregated residual transformations for deep neural networks. **Proceedings of the IEEE conference on computer vision and pattern recognition**, IEEE, p. 1492–1500, 2017.

YUE, Y.; LIU, D.; FU, S.; ZHOU, X. Heart rate and heart rate variability as classification features for mental fatigue using short-term ppg signals via smartphones instead of ecg recordings. **2021 13th International Conference on Communication Software and Networks (ICCSN)**, IEEE, p. 370–376, 2021.

ZAGORUYKO, S.; KOMODAKIS, N. Wide residual networks. **arXiv preprint**, Cornell University, Nova York, v. 4, p. 282 – 292, 2016.

ZHANG, H. The optimality of naive bayes. **University of New Brunswick**, v. 1, n. 2, p. 3, 2004.

ZHANG, X.; ZHOU, X.; LIN, M.; SUN, J. Shufflenet: An extremely efficient convolutional neural network for mobile devices. **Proceedings of the IEEE conference on computer vision and pattern recognition**, IEEE, p. 6848–6856, 2018.