



Exploração de Técnicas de *Clustering* Aplicadas a Patologias da Voz

ALESSA ANJOS DE OLIVEIRA

Dissertação para obtenção do grau de Mestre em:

Engenharia Industrial – Ramo Engenharia Eletrotécnica

Este trabalho foi efetuado sob orientação de:

Prof. Dr. João Paulo Ramos Teixeira

Prof. Dra. Maria Eugênia Dajer

Bragança

Junho de 2020



ALESSA ANJOS DE OLIVEIRA

Exploração de Técnicas de *Clustering* Aplicadas a Patologias da Voz

Dissertação apresentada como requisito do acordo de Dupla Diplomação entre a UTFPR e o IPB para a obtenção do título de Mestre em Engenharia Industrial – Ramo Engenharia Eletrotécnica

Orientadores:

Prof. Dr. João Paulo Ramos Teixeira

Prof. Dra. Maria Eugênia Dajer

Bragança

19 de Junho de 2020

Agradecimentos

Em primeiro lugar gostaria de agradecer à minha família por todo o suporte e exemplos que me proporcionaram. A minha mãe, Elizabeth M. A. Oliveira, por todo o carinho, dedicação e conselhos. Ao meu pai, Edison de Oliveira, por me ensinar a ter paciência e me apoiar em todas as escolhas. Aos meus irmãos, Alison Anjos de Oliveira e Andressa Anjos de Oliveira, por me moldarem em vários aspectos da minha personalidade e me ensinarem a me posicionar diante a adversidades. Aos meus avós, Arthur M. dos Anjos e Ilza Escolar dos Anjos por me ensinarem muito sobre resiliência. Agradeço também à minha bisavó, mais carinhosamente conhecida por nona, por todo o amor e histórias de vida.

Meu muito obrigada ao Iuri M. S. M. de Carvalho por todo amparo e amor nesse trajeto, mesmo em dias difíceis. Agradeço também aos amigos feitos por todo esse caminho percorrido, me proporcionando momentos de alegria e diversão no meio de tanto caos. Nomeadamente, Fernanda Cristina Sartori da Silva, Mariane Beatriz de Souza Correia, Jean Rafael da Silva, Thiago Castro Bueno, Gustavo Monteiro Romero, Geovana Vilharquide, Carlos César Egídio, Miliani Silva Timofó Ferreira, Lucas Fernandes de Azevedo, Chrysologo Rocha de Oliveira Neto, Bárbara Giehl e Carolina Magalhães Cezar.

Agradeço aos professores que, ao longo do meu trajeto acadêmico, me apoiaram e acreditaram no meu potencial de poder sempre melhorar. Nominalmente os Drs. Marcos Vallim, Emerson Ravazzi, Luiz Buzachero e Carolina Rodrigues.

Meu principal agradecimento é destinado ao Instituto Politécnico de Bragança e a Universidade Tecnológica Federal do Paraná campus Cornélio Procópio por me proporcionarem essa oportunidade única. Amplio minha mais sincera gratidão aos professores Drs. João Paulo Teixeira e Maria Eugênia Dajer pela orientação e a concessão de sabedoria.

“Nunca estou realmente satisfeita quanto a entender alguma coisa; porque, até onde entendo, a minha compreensão só pode ser uma fração infinitesimal de tudo o que eu quero compreender.”

Ada Lovelace

Resumo

Este trabalho de dissertação de mestrado tem o intuito de aplicar técnicas de *clustering* em um conjunto de dados a fim de verificar a possibilidade de agrupamento entre grupos que contém informações a respeito da voz. Neste estudo são empregados os parâmetros acústicos *jitter* absoluto, *jitter* relativo, *shimmer* absoluto, *shimmer* relativo, autocorrelação, *Harmonic to Noise Ratio* e *Noise to Harmonic Ratio*. Estes atributos são extraídos da base de dados *Saarbrücken Voice Database*, a partir de ficheiros de áudio que dispõem das vogais /a/, /i/ e /u/ sustentadas nos tons alto, baixo e normal. Para o agrupamento dos conjuntos através desses parâmetros são utilizados o grupo de Controlo, o Carcinoma, o Cisto, a Disfonia Espasmódica, a Disfonia Funcional, a Disfonia Hiperfuncional, a Disfonia Hipofuncional, a Disfonia Hipotônica, a Disfonia Psicogênica, o Edema de Reinke, o Granuloma, o Granuloma por Intubação, a Laringite Crônica, a Paralisia das Cordas Vocais, Pólipos das Cordas Vocais, o Tumor na Hipofaringe e o Tumor na Laringe. Esta investigação ocorre ao empregar 2 caminhos diferentes para averiguar a possibilidade de agrupar patologias: a análise estatística descritiva e algoritmos de agrupamento. Para a análise estatística descritiva, utiliza-se o Diagrama de Caixas. Para os métodos de agrupamento, são empregados os Mapas Auto-Organizáveis de Kohonen e o *clustering* hierárquico aglomerativo. Os resultados dessas 2 categorias apontam a divisão dos dados em 2 conjuntos, em que um grupo contém grande parte das amostras, enquanto o outro contém apenas algumas amostras. No entanto, essas técnicas se divergem ao comparar os elementos pertencentes aos conjuntos, pois, na análise estatística descritiva as amostras se encontram mais bem separadas do que nos métodos de *clustering*.

Palavras Chave: Análise Estatística Descritiva, Mapas Auto-Organizáveis de Kohonen, *Clustering* Hierárquico, Técnicas de *Clustering*, Parâmetros Acústicos, Diagrama de Caixa.

Abstract

This dissertation work aims to apply *clustering* techniques to a set of data in order to verify the possibility of *clustering* between groups that contain information about the voice. In this study, the acoustic parameters absolute *jitter*, relative *jitter*, absolute *shimmer*, relative *shimmer*, autocorrelation, *Harmonic to Noise Ratio* and *Noise to Harmonic Ratio* are used. These attributes are extracted from the *Saarbrücken Voice Database*, from audio files that have the vowels /a/, /i/ and /u/ sustained in a high, low and normal tones. For the *clustering* of the sets using these parameters, the Control group, Carcinoma, Cyst, Spasmodic Dysphonia, Functional Dysphonia, Hyperfunctional Dysphonia, Hypofunctional Dysphonia, Hypotonic Dysphonia, Psychogenic Dysphonia, Reinke's Edema, Granuloma, Intubation Granuloma, Chronic Laryngitis, Vocal Cord Paralysis, Vocal Cord Polyps, Hypopharyngeal Tumor and Laryngeal Tumor are utilized. This investigation occurs when employing 2 different ways to investigate the possibility of grouping pathologies: the descriptive statistical analysis and *clustering* algorithms. For the descriptive statistical analysis, the Box Diagram is used. Self-Organizing Maps of Kohonen and agglomerative hierarchical *clustering* are utilized for grouping methods. The results of these 2 categories point to the division of the data into 2 sets, in which one group contains a large part of the samples, while the other contains only a few samples. However, these techniques differ when comparing the elements belonging to the sets, because, in the descriptive statistical analysis, the samples are better divided than in the *clustering* methods.

Keywords: Descriptive Statistical Analysis, Self-Organizing Maps of Kohonen, Hierarchical *Clustering*, *Clustering* Techniques, Acoustic Parameters, Box Diagram.

Índice de Figuras

Figura 1 - (a) Paralisia unilateral incompleta; (b) paralisia bilateral incompleta; (c) paralisia unilateral completa e (d) paralisia bilateral completa. Fonte: Fawcus (1991).....	11
Figura 2 – Representação gráfica da construção de um Diagrama de Caixa.	18
Figura 3 – Primeira situação: há uma diferença entre os grupos. Fonte: Teixeira et al. (2018).....	19
Figura 4 – Segunda situação: há, provavelmente, uma diferença entre os grupos. Fonte: Teixeira et al. (2018).	19
Figura 5 – Terceira situação: não há diferença entre os grupos. Fonte: Teixeira et al. (2018).	19
Figura 6 – Diagrama de Caixa comparando os conjuntos da base de dados por meio do <i>jitter</i> absoluto para o gênero feminino.	21
Figura 7 – Diagrama de Caixa comparando os conjuntos da base de dados por meio do <i>jitter</i> absoluto para o gênero masculino.	23
Figura 8 – Diagrama de Caixa do <i>jitter</i> relativo feito entre os conjuntos de teste.	24
Figura 9 – Diagrama de Caixa comparando os conjuntos extraídos do SVD por meio do <i>shimmer</i> absoluto.	26
Figura 10 – Diagrama de Caixa comparando os conjuntos por meio do <i>shimmer</i> relativo.	27
Figura 11 – Diagrama de Caixa utilizando o HNR para analisar os conjuntos do banco de dados.	29
Figura 12 – Comparação dos conjuntos através do Diagrama de Caixa utilizando o NHR.	30
Figura 13 - Utilização da autocorrelação no estudo do Diagrama de Caixa para os grupos de teste.	31
Figura 14 - Diagrama de métodos de <i>clustering</i> convencionais. Adaptado de Gan (2007).	36
Figura 15 – Métodos de agrupamento hierárquico. Adaptado de Gan (2007).....	38
Figura 16 – Dendrograma de 5 elementos. Fonte: Gan (2007).....	39
Figura 17 - Distância entre a vizinhança dos neurônios.	43
Figura 18 – Árvore hierárquica aglomerativa dividida em 2 conjuntos.	44

Figura 19 – Distância entre a vizinhança dos neurônios com o novo conjunto.	45
Figura 20 - Árvore hierárquica aglomerativa com novo agrupamento.....	46

Índice de Tabelas

<i>Tabela 1 – Definição dos dados extraídos do SVD.....</i>	<i>6</i>
<i>Tabela 2 – Resultado das semelhanças entre os conjuntos com base no jitter absoluto feminino.....</i>	<i>22</i>
<i>Tabela 3 – Semelhança entre os conjuntos com base no Diagrama de Caixa.</i>	<i>23</i>
<i>Tabela 4 – Resultados das comparações por meio do Diagrama de Caixa para o jitter relativo.....</i>	<i>25</i>
<i>Tabela 5 – Resultado da comparação entre as caixas do parâmetro shimmer absoluto.</i>	<i>26</i>
<i>Tabela 6 – Possíveis agrupamentos por meio do Diagrama de Caixa do shimmer relativo. ...</i>	<i>28</i>
<i>Tabela 7 – Resultado do Diagrama de Caixa utilizando as amostras do HNR sobre semelhanças entre os grupos.</i>	<i>29</i>
<i>Tabela 8 – Resultados das comparações por Diagrama de Caixa dos conjuntos de teste através do parâmetro NHR.....</i>	<i>30</i>
<i>Tabela 9 – Resultado das semelhanças entre os grupos de teste por meio da autocorrelação.....</i>	<i>32</i>
<i>Tabela 10 - Relações entre os grupos de teste com base na comparação de Diagrama de Caixas.....</i>	<i>33</i>
<i>Tabela 11 – Relação entre métricas de ligação e suas distâncias copenéticas.....</i>	<i>42</i>

Acrónimos

GMM	<i>Gaussian Mixture Models</i>
HNR	<i>Harmonic to Noise Ratio</i>
IQR	Intervalo Interquartil
<i>jitta</i>	<i>Jitter Absoluto</i>
<i>jitter</i>	<i>Jitter Relativo</i>
KNN	<i>K-Nearest Neighbors</i>
LPCC	<i>Linear Prediction Cepstral Coefficient</i>
MDVP	<i>Multi-Dimensional Voice Program</i>
MFCC	<i>Mel-Frequency Cepstral Coefficient</i>
ML	<i>Machine Learning</i>
NHR	<i>Noise to Harmonic Ratio</i>
PLP	<i>Perceptual Linear Prediction</i>
RNA	Rede Neuronal Artificial
ShdB	<i>Shimmer Absoluto</i>
Shim	<i>Shimmer Relativo</i>
SOM	Mapa Auto Organizável de Kohonen
SVD	<i>Saarbruecken Voice Database</i>

Índice

CAPÍTULO 1: INTRODUÇÃO	1
1.1. CONTEXTUALIZAÇÃO	1
1.2. MOTIVAÇÃO E JUSTIFICATIVA	2
1.3. OBJETIVOS	3
1.4. ESTADO DA ARTE	3
1.5. ORGANIZAÇÃO DO TRABALHO	4
CAPÍTULO 2: A VOZ	5
2.1. CONTEXTUALIZAÇÃO	5
2.2. BASE DE DADOS	6
2.3. DESCRIÇÃO DOS GRUPOS DE TESTE	7
2.3.1. <i>Voz Saudável</i>	7
2.3.2. <i>Carcinoma</i>	7
2.3.3. <i>Cisto</i>	8
2.3.4. <i>Disfonia Espasmódica</i>	8
2.3.5. <i>Disfonia Funcional</i>	8
2.3.6. <i>Disfonia Hiperfuncional</i>	8
2.3.7. <i>Disfonia Hipofuncional</i>	9
2.3.8. <i>Disfonia Hipotônica</i>	9
2.3.9. <i>Disfonia Psicogênica</i>	9
2.3.10. <i>Edema de Reinke</i>	10
2.3.11. <i>Granuloma</i>	10
2.3.12. <i>Granuloma por Intubação</i>	10
2.3.13. <i>Laringite Crônica</i>	10
2.3.14. <i>Paralisia das Cordas Vocais</i>	11
2.3.15. <i>Pólipos das Cordas Vocais</i>	11
2.3.16. <i>Tumor na Hipofaringe</i>	12
2.3.17. <i>Tumor na Laringe</i>	12
2.4. PARÂMETROS DOS SINAIS ACÚSTICOS	12
2.4.1. PERTURBAÇÕES EM FREQUÊNCIA	13
2.4.2. <i>Perturbações em amplitude</i>	13
2.4.3. <i>Autocorrelação</i>	14
2.4.4. <i>HNR</i>	15
2.4.5. <i>NHR</i>	16

CAPÍTULO 3: ANÁLISE ESTATÍSTICA DESCRITIVA.....	17
3.1. CONTEXTUALIZAÇÃO.....	17
3.2. <i>Diagrama de Caixa</i>	17
3.3. <i>Conferência entre Pares de Caixas</i>	19
3.4. METODOLOGIA.....	20
3.5. RESULTADOS E DISCUSSÕES	20
3.5.1. <i>JITTER ABSOLUTO FEMININO</i>	21
3.5.2. <i>JITTER ABSOLUTO MASCULINO</i>	22
3.5.3. <i>JITTER RELATIVO</i>	24
3.5.4. <i>SHIMMER ABSOLUTO</i>	25
3.5.5. <i>SHIMMER RELATIVO</i>	27
3.5.6. <i>HNR</i>	28
3.5.7. <i>NHR</i>	30
3.5.8. <i>AUTOCORRELAÇÃO</i>	31
3.6. CONCLUSÃO.....	33
CAPÍTULO 4: MÉTODOS DE CLUSTERING	35
4.1. CONTEXTUALIZAÇÃO.....	35
4.2. MAPAS AUTO-ORGANIZÁVEIS DE KOHONEN	36
4.3. <i>CLUSTERING</i> HIERÁRQUICO.....	38
CAPÍTULO 5: DESENVOLVIMENTO PRÁTICO DOS MÉTODOS DE AGRUPAMENTO	40
5.1. METODOLOGIA.....	40
5.2. MAPAS AUTO-ORGANIZÁVEIS DE KOHONEN	41
5.3. <i>CLUSTERING</i> HIERÁRQUICO AGLOMERATIVO	41
5.4. RESULTADOS E DISCUSSÕES	42
5.5. CONCLUSÃO.....	47
CAPÍTULO 6: CONCLUSÃO E TRABALHOS FUTUROS	49
6.1. CONCLUSÃO GERAL.....	49
6.2. TRABALHOS FUTUROS.....	51
REFERÊNCIAS	52

Capítulo 1: Introdução

Este primeiro capítulo aborda a introdução do tema, seguido pelos objetivos dessa dissertação. Além de designar o Estado da Arte, que possui trabalhos relacionados a esse estudo. Por fim, é disposta a estrutura desta dissertação.

1.1. Contextualização

A voz é o método mais importante da comunicação humana. E cada ser humano possui uma voz distinta (Rosa, Guimarães & Caçador, 2011). A voz é produzida por meio do ar expirado pelo pulmão, fazendo com que ocorra a vibração das pregas vocais. Essa vibração percorre um longo caminho, sofrendo alterações acústicas até alcançar som da voz humana (Rosa et al., 2011). Esse sinal pode sofrer diversas alterações, devido à ingestão de produtos tóxicos, uso vocal incorreto ou problemas hormonais e neurológicos (Cordeiro, 2016), desenvolvendo, assim, possíveis disfunções. Além do mais, a qualidade da voz também pode se alterar em consequência do envelhecimento e amadurecimento dos indivíduos (Soltani, Ashayeri, Modarresi, Salavati & Ghomashchi, 2014).

Para investigar esses distúrbios que afetam a qualidade vocal pode ser feita uma análise perceptiva auditiva, contudo, essa é uma técnica subjetiva que pode apresentar diferentes resultados conforme a experiência médica do profissional (Teixeira, Oliveira & Lopes, 2013), (Teixeira & Fernandes, 2014) ou podem ser invasivas (Teixeira, Ferreira & Carneiro, 2011). Uma análise mais objetiva pode ser feita através de parâmetros acústicos extraídos da voz por meio do processamento desse sinal de fala (Teixeira et al., 2013), (Teixeira & Fernandes, 2014). Com esses atributos é possível comparar sujeitos que possuem vozes consideradas saudáveis com pessoas que dispõem de alguma patologia.

Pelo fato de a voz ser um sinal complexo, sua qualidade pode ser discutida através de vários parâmetros do sinal acústico. Através desses mesmos atributos é possível investigar como funciona a estrutura desses dados, aplicando métodos de análise exploratória.

A área de *clustering*, na análise exploratória de dados, organiza os objetos de estudo com base em suas similaridades, estruturando-os por métodos particionais ou hierárquicos (Jain & Dubes, 1988). No entanto, esses métodos utilizam de aprendizado não-supervisionado, onde

não há informações a respeito das saídas desejadas, sendo necessária a obtenção de informações complementares a respeito dos dados estudados (Silva, Spatti & Flauzino, 2016). Para tanto, empregar a análise estatística descritiva e comparar os dados com informações a respeito da qualidade vocal podem aprimorar o conhecimento a respeito dos grupos resultantes apresentados neste estudo, além de fornecer maior noção a respeito da estrutura dos dados.

O Mapa Auto Organizável de Kohonen (SOM) é uma técnica de *clustering* por partição capaz de reproduzir um conjunto multidimensional em bidimensional (Affonso, 2011), (Gan, Ma & Wu, 2007), (Kohonen, 1995), (Silva et al., 2016), apresentando uma boa visualização do agrupamento de dados. O *clustering* hierárquico aglomerativo aponta informações quantitativas a respeito desses agrupamentos. Portanto, a investigação utilizando 2 Redes Neurais Artificiais (RNAs) distintas pode fornecer noções consideráveis a respeito do sinal de voz, por serem métodos divergentes capazes de fornecer informações que complementam os resultados apresentados por ambas RNAs.

1.2. Motivação e Justificativa

Técnicas tradicionais de detecção de patologias da voz podem ser invasivas ou subjetivas, como exames laringoscópios ou percepção auditiva, respectivamente (Teixeira et al., 2011). O processamento do sinal de fala é capaz de extrair vários parâmetros do sinal acústico capazes de avaliar a qualidade da voz de indivíduos (Teixeira et al., 2013), (Teixeira & Fernandes, 2014). De acordo com Rosa et al. (2011), a voz é um fenômeno multidimensional, sendo um dos recursos mais complexos dos seres humanos. Conseqüentemente, essa característica pode trazer dificuldades para classificar as vozes em patológicas ou saudáveis.

Para tanto, faz-se uso de ferramentas poderosas, como as RNAs para averiguar se os atributos extraídos do banco de dados são capazes de se agrupar. O estudo de agrupamentos é uma área da exploração de dados que vem crescendo abruptamente (Jain & Dubes, 1988), podendo ser empregada em diversos campos distintos, como engenharia, medicina, biologia e *marketing* (Jain & Dubes, 1988).

Este trabalho faz uso de técnicas de *clustering* para agrupar patologias com base nos parâmetros disponíveis na base de dados investigada, a fim de aumentar a dimensão de grupos da base de dados para usar em classificadores inteligentes.

1.3. Objetivos

Essa dissertação tem como objetivo aplicar técnicas de *clustering* em sinais de voz de uma base de dados tratada a fim de averiguar se é possível o agrupamento desses dados com base nos parâmetros aferidos. Tais ferramentas tem o propósito de compreender como os atributos da voz se estruturam para agrupar-se em *clusters*.

Nomeadamente, considera-se os seguintes objetivos:

- Aplicação de uma análise estatística descritiva;
- Implementação de algoritmos de *clustering*, mais especificamente:
 - Mapas Auto-Organizáveis de Kohonen;
 - Agrupamento hierárquico aglomerativo.

1.4. Estado da Arte

Esta seção traz uma revisão a respeito de investigações relacionadas a métodos de *clustering* através parâmetros acústicos retirados de sinais vocais. As técnicas de agrupamento deste estudo não foram encontradas em conjunto como aplica-se nessa dissertação, portanto, esses métodos foram pesquisados separadamente e organizados em ordem cronológica.

Em 1966, Herman R. Silbiger fez uma apresentação de 12 minutos dissertando a respeito do *clustering* hierárquico para classificar vozes de 24 falantes do sexo masculino e 23 do sexo feminino por meio das distâncias entre os perfis diferenciais semânticos desses indivíduos. O estudo demonstrou ser um método promissor para a seleção de uma amostra de falantes com vozes perceptivamente diferentes.

Leinonen, Kangas, Torkkola, Juvas publicaram, em 1992, um artigo sobre a utilização de Mapas Auto-Organizáveis para encontrar índices a respeito as qualidades da voz disfônica. Foram extraídos espectros de potência a curto prazo de 24 pacientes e 58 sujeitos sem problemas vocais que pronunciavam palavras finlandesas em uma sonoridade confortável. Por meio do mapa de Kohonen verificou-se a existência de diferenças distintas entre as amostras consideradas normais e aquelas consideradas disfônicas, revelando diferenças estatisticamente significativas entre as vozes normal e disfônica.

Em 1994, Rihkanen, Leinonen, Hiltunen e Kangas aplicaram a rede de Kohonen, dentre tantas medidas de voz, em espectros de frequências. Eles tinham como objetivo descobrir se

esse método distinguiria entre vozes saudáveis e disfônicas e se detectaria alterações devido à cirurgia da laringe. Para tal, 49 amostras da vogal /a/ foram coletadas de 19 indivíduos sem problemas de voz, 15 sujeitos antes e após a modificação da estrutura da cartilagem tireóidea de uma corda vocal paralisada ou atrofiada. O mapa final diferenciou vozes normais e disfônicas com base na composição espectral de maneira estatisticamente significativa.

Leinonen, Hiltunen, Laakso, Rihkanen e Poppins investigaram o nível de patologia, aspereza, soprosidade, tensão, astenia e *pitch* para categorizar vozes em 1997. Por meio de uma lista de palavras, 58 homens e 82 mulheres foram analisados. Alguns sujeitos apresentaram uma grande variedade de diagnósticos, mas a maioria deles não havia sido tratado para nenhum distúrbio da voz. De acordo com os resultados dispostos pelo SOM, as qualidades da voz disfônica foram descritas como aglomerados, sendo possível estabelecer uma base para a avaliação acústica objetiva da qualidade da voz.

1.5. Organização do Trabalho

Esse trabalho se organiza em 6 tópicos. O Capítulo 1 contextualiza o tema dessa dissertação, sendo abordados a motivação e justificativa, objetivos e Estado da Arte. No Capítulo 2 são definidas informações a respeito da voz, tais quais a base de dados empregada, a descrição dos grupos investigados e os parâmetros acústicos. O Capítulo 3 contém a fundamentação teórica e desenvolvimento prático da análise estatística descritiva. O Capítulo 4 emprega a fundamentação teórica a respeito dos métodos de *clustering* investigados nesse estudo. No Capítulo 5 é apresentado o desenvolvimento prático das técnicas de agrupamento. O Capítulo 6, expõe a conclusão geral desse estudo e inclui sugestões para trabalhos futuros.

Capítulo 2: A Voz

Neste capítulo é exposto conceitos a respeito da voz, a base de dados empregada, a descrição dos conjuntos analisados e os parâmetros acústicos utilizados neste trabalho.

2.1. Contextualização

O indivíduo saudável, para Cordeiro (2016), é capaz de produzir vibrações nas pregas vocais com uma intensidade quase constante e de forma periódica. Todavia, mesmo uma pessoa saudável já passou por uma mudança na voz, devido a algum distúrbio vocal, seja ele um resfriado, uma alergia ou até mesmo algo mais sério (Boone, McFarlane, Von Berg & Zraick, 2014). A voz pode ser alterada em virtude das condições físicas e psicológicas em que o indivíduo se encontra, contudo, é perceptível que “quando há alterações patológicas na laringe, o nível do sinal e a sua frequência fundamental mudam” (Panek, Skalski, Gajda & Tadeusiewicz, 2015), sendo assim, essa variação sonora pode ser audível.

De maneira genérica, as patologias da voz podem se ramificar em 3 conjuntos: funcional, orgânico e neurogênico (Boone et al., 2014). O primeiro distúrbio não afeta a estrutura do aparelho vocal. A segunda categoria é relacionada a alterações na estrutura do trato vocal ou a doenças que possam surgir em partes específicas desse local. O terceiro grupo abrange distúrbios que decorrem de problemas no sistema nervoso central ou periférico.

Com o processamento de sinais é possível extrair um conjunto de parâmetros que colaboram para a detecção de distúrbios na laringe. Para isso, é necessário que haja um sinal gravado de forma ininterrupta e constante tanto de indivíduos com uma voz considerada saudável quanto pacientes que possuem algum distúrbio na voz. Assim, é possível retirar os parâmetros *jitter* absoluto, *jitter* relativo, *shimmer* absoluto, *shimmer* relativo, *Harmonic to Noise Ratio* (HNR), *Noise to Harmonic Ratio* (NHR) e autocorrelação dos sinais de voz.

2.2. Base de Dados

O *Saarbruecken Voice Database* (SVD) é um acervo de vozes online de forma aberta disponibilizado pelo Instituto de Fonética da Universidade de Saarland. (Pützer & Barry, s.d.). Esta coleção contém gravações de mais de 2000 pessoas de cidadania alemã com 71 patologias diferentes mais o grupo de Controlo (Pützer & Barry, s.d.).

No SVD estão dispostas gravações das vogais sustentadas /a/, /i/ e /u/ produzidas em tom normal, alto, baixo e com variação entre tons. Além da pronuncia da frase "*Guten Morgen, wie geht es Ihnen?*" ("Bom dia, como vai você?"), totalizando 13 arquivos por pessoa. Nesta base de dados todas as gravações apresentam uma frequência de amostragem de 50 kHz e resolução de 16 bits (Pützer & Barry, s.d.).

Nesta dissertação são utilizadas as vogais sustentadas /a/, /i/ e /u/ nos tons alto, baixo e normal dos grupos de Controlo, Carcinoma, Cisto, Disfonia Espasmódica, Disfonia Funcional, Disfonia Hiperfuncional, Disfonia Hipofuncional, Disfonia Hipotônica, Disfonia Psicogênica, Edema de Reinke, Granuloma, Granuloma por Intubação, Laringite Crônica, Paralisia das Cordas Vocais, Pólipos das Cordas Vocais, Tumor na Hipofaringe e Tumor na Laringe. A Tabela 1 apresenta o tamanho das amostras a serem utilizadas neste trabalho, obtendo um total de 830 indivíduos. Como cada indivíduo tem um ficheiro com as 3 vogais sendo pronunciadas em 3 tons diferentes, tem-se 9 variáveis para cada indivíduo, logo, a base de dados tem tamanho de 7470 amostras.

Tabela 1 – Definição dos dados extraídos do SVD.

Grupos de Teste	Idade	Tamanho da Amostra		
	Média	Feminino	Masculino	Total
Controle	36,74	123	71	194
Carcinoma	62,06	1	18	19
Cisto	42,00	2	1	3
Disfonia Espasmódica	53,75	40	22	62
Disfonia Funcional	46,89	51	24	75
Disfonia Hiperfuncional	40,91	95	32	127
Disfonia Hipofuncional	45,83	4	8	12
Disfonia Hipotônica	49,50	0	2	2
Disfonia Psicogênica	51,29	38	13	51
Edema de Reinke	54,47	29	5	34
Granuloma	44,50	1	1	2
Granuloma por Intubação	53,00	0	3	3
Laringite Crônica	49,80	16	25	41
Paralisia nas Cordas Vocais	57,52	102	67	169

Pólipo nas Cordas Vocais	52,04	10	17	27
Tumor na Hipofaringe	57,20	0	5	5
Tumor na Laringe	49,00	1	3	4
Total				830

2.3. Descrição dos Grupos de Teste

Neste tópico serão discutidos os 17 conjuntos escolhidos para a análise acústica, sendo eles o grupo de Controlo, enquadrando-se na voz saudável, o Carcinoma, o Cisto, a Disfonia Espasmódica, a Disfonia Funcional, a Disfonia Hiperfuncional, a Disfonia Hipofuncional, a Disfonia Hipotônica, a Disfonia Psicogênica, o Edema de Reinke, o Granuloma, o Granuloma por Intubação, a Laringite Crônica, a Paralisia das Cordas Vocais, o Pólipo das Cordas Vocais, o Tumor na Hipofaringe e o Tumor na Laringe.

2.3.1. Voz Saudável

A voz saudável é fruto da vibração eficiente das pregas vocais. Essa eficiência advém da sincronização da contração muscular adequada e uma estrutura respiratória harmoniosa. Caso haja alteração em algum dos fatores supracitados, a qualidade da voz produzida é alterada. (LaPine, 2008), (Medscape, s.d.), (NIDCD, s.d.).

Pontes, Yamasaki & Behlau (2006) apontam que jovens que apresentam vozes normais exprimem o fechamento glótico completo, contudo, alguns fatores, como diferenças culturais, sociais e genéticas, podem afetar essa ação de fechamento. Ademais, a qualidade da voz se altera devido ao tamanho e formato das pregas vocais de cada indivíduo, além de razões como maturidade, envelhecimento e variação na precisão do controlo da fala (Soltani et al., 2014). Por conseguinte, há uma discrepância no que pode ser definido como características que definem uma boa qualidade vocal.

2.3.2. Carcinoma

O Carcinoma se desenvolve quando pequenas áreas de células anômalas sofrem alterações sequenciais, crescendo fora de controle até evoluir para um câncer. Tem como

gatilho o tabagismo e consumo excessivo de álcool. O principal sintoma deste distúrbio é a rouquidão, apesar de haverem outros indícios, tais quais dor crônica de garganta, dificuldade em respirar e engolir, perda de peso não intencional e existência de nódulos no pescoço (John Hopkins Medicine, s.d.).

2.3.3. Cisto

O Cisto na laringe pode ser ramificado em dois tipos: retenção de muco, onde há uma obstrução anormal no ducto glandular, havendo a retenção de secreções; e sebáceo, que é fruto de falhas durante o processo de cura da mucosa lesionada sobrepondo a pele ou no desenvolvimento embriológico (Altman, 2007), (Cleveland Clinic, s.d.). Podem ser ocasionados devido ao uso excessivo da voz quando há uma infecção respiratória e tem como sintoma a rouquidão, ocasionada por anomalias nas cordas vocais.

2.3.4. Disfonia Espasmódica

A Disfonia Espasmódica é um distúrbio neurológico que atinge os músculos das cordas vocais, causando movimentos involuntários abruptos que impedem a produção efetiva da voz, onde a pessoa exibe dificuldade na produção de uma ou mais palavras (NIDCD, s.d.). Boone et al. (2014) apontam que o indivíduo com essa disfunção “apresenta uma voz estridente e estrangulada, com esforço observável para empurrar o ar para fora durante a maioria das tentativas de voz”.

2.3.5. Disfonia Funcional

Na Disfonia Funcional não há alteração anatômica nas pregas vocais, logo, o distúrbio é ocasionado pelo mau uso da voz ou por alterações psicológicas. Essa patologia tem como sintomas voz ofegante, fadiga e instabilidade da voz (Michigan Medicine, s.d.). Assim, a qualidade, tom e intensidade da voz são irregulares, mesmo a fonação se mantendo (Roy, 2003).

2.3.6. Disfonia Hiperfuncional

A Disfonia Hiperfuncional é uma das ramificações da Disfonia Funcional que causa o aumento de tensão no músculo laríngeo, resultando numa rigidez excessiva das pregas vocais,

aproximando-as, podendo ser acompanhada pela ampliação da força do ar subglotal e amplitudes mais elevadas das pregas vocais (Holmberg, Doyle, Perkell, Hammarberg & Hillman, 2002), (Kandoğan, Koç & Aksoy, 2009), (Hillman, Holmberg, Perkell, Walsh & Vaughan, 1989). A voz resultante dessa patologia é estridente e, ao mesmo tempo, fadigada. Com o acréscimo de componentes aperiódicos ao sinal acústico, a voz pode aparentar rouquidão (Hillman et al., 1989).

2.3.7. Disfonia Hipofuncional

A Disfonia Hipofuncional é, também, uma parcela da Disfonia Funcional que, ao contrário da Disfonia Hiperfuncional, apresenta o fechamento incompleto das cordas vocais devido a uma tensão muscular frágil na fonação. “Esse distúrbio de voz é caracterizado principalmente por fonação respiratória, grave e silenciosa” (Dr. Wohlt, s.d.).

2.3.8. Disfonia Hipotônica

A Disfonia Hipotônica é uma Disfonia Funcional que apresenta o enfraquecimento da musculatura da laringe, impedindo o fechamento completo das pregas vocais, o que faz com que o paciente apresente uma voz rouca e fraca (Clínica San Vicente, s.d.). Essa patologia é associada a defeitos nas estruturas do aparelho vocal e problemas congênitos.

2.3.9. Disfonia Psicogênica

A Disfonia Psicogênica é causada por uma somatização de processos psicológicos negativos, como ansiedade, estresse, depressão e conflitos no ambiente de trabalho ou profissional, que resultam na mudança brusca da qualidade da voz (Martins, Tavares, Ranalli, Branco & Pessin, 2014). Martins et al. (2014) explicam que “o controle respiratório, intensidade vocal, alcance vocal, ressonância vocal, frequência fundamental, articulação e velocidade e entonação da fala podem estar comprometidos. Na maioria dos casos, mais de um parâmetro vocal é alterado, permanentemente ou não”.

2.3.10. Edema de Reinke

O Edema de Reinke ocorre quando o espaço de Reinke se enche de líquidos devido a traumas vocais de longa data, como o tabagismo, abuso vocal recorrente e refluxo gástrico (Columbia University, s.d.). “O edema de Reinke produz uma voz profunda e rouca” (Altman, 2007). É visto que mulheres mais velhas com histórico de tabagismo possuem maior chances de desenvolver esse distúrbio.

2.3.11. Granuloma

O Granuloma é uma lesão benigna resultante de um processo reativo a um ferimento no processo vocal (Morrison, Nichol & Rammage, 1994). Os principais fatores que levam a essa condição são uso inadequado da voz, refluxo de fluidos estomacais e irritação devido ao uso de tubo respiratório.

2.3.12. Granuloma por Intubação

O Granuloma por Intubação se dá quando o paciente é intubado por longos períodos de tempo, fazendo com que haja atrito na parte posterior da laringe, desenvolvendo uma lesão. Assim, o tecido inflamatório da granulação se forma e, caso haja, um impedimento da cicatrização pode vir a se tornar um tumor localizado (British Medical Journal, 1973). Segundo Rimoli, Martins, Catâneo, Imamura & Catâneo (2018) “os tratamentos recomendados para granulomas por intubação incluem terapia de voz, medicamentos anti-refluxo, cirurgia, toxina butolímica, sulfato de zinco, mitomicina C e esteroides”.

2.3.13. Laringite Crônica

A inflamação nas pregas vocais é chamada de Laringite. Essa disfunção deturpa os sons produzidos pelas cordas vocais, deixando a voz rouca. A Laringite Crônica tem duração superior a três semanas e é causada por exposição de substâncias irritantes ou uso excessivo da voz por um longo período de tempo (Mayo Clinic, s.d.).

2.3.14. Paralisia das Cordas Vocais

A Paralisia das Cordas Vocais é sintoma de uma lesão no nervo laríngeo recorrente, incapacitando a contração muscular das pregas vocais, que pode ser ocasionada por trauma, neoplasia, cirurgia ou idiopatia (Chen, Jen, Wang, Lee & Lin, 2007), (Holinger, Holinger & Holinger, 1976), (Parnell & Brandenburg, 1970); (Toutounchi, Eydi, Golzari, Ghaffari & Parvizian, 2014). Esse distúrbio, conforme exposto por Fawcus (1991), pode se decompor em quatro situações: paralisia unilateral incompleta, paralisia bilateral incompleta, paralisia unilateral completa e paralisia bilateral completa. A Figura 1 apresenta os quatro casos de paralisia.

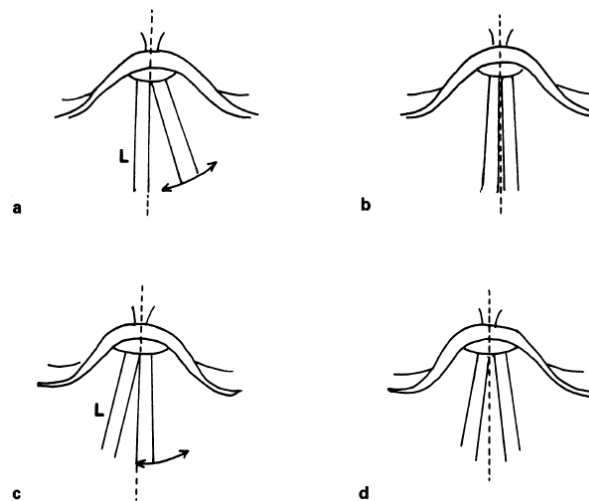


Figura 1 - (a) Paralisia unilateral incompleta; (b) paralisia bilateral incompleta; (c) paralisia unilateral completa e (d) paralisia bilateral completa. Fonte: Fawcus (1991).

De acordo com Holinger et al. (1976), a paralisia bilateral pode soar suficientemente normal, pelo fato de que ambas pregas permanecerem imobilizados em adução, onde vibram fracamente, mas iguais. No caso da paralisia unilateral, o paciente apresenta uma voz rouca e aspirada, devido as diferenças de tensão entre as cordas normal e paralisada.

2.3.15. Pólipos das Cordas Vocais

O Pólipo nas Cordas Vocais pode ser dito como uma anormalidade benigna que ocorre em resposta ao rompimento de um pequeno vaso sanguíneo, onde o corpo tenta curar essa lesão formando uma camada de tecido que se transforma no pólipo (University of Miami Health

System, s.d.). Esse distúrbio decorre de processos irritantes para as pregas vocais, como o uso contínuo da voz em um tom elevado, tabagismo, aspiração de substâncias químicas agressivas, uso de álcool e alergias, afetando a qualidade da voz. Os principais sintomas desta disfunção são a rouquidão, fadiga vocal, pigarro e sensação de nó na garganta.

2.3.16. Tumor na Hipofaringe

O Tumor na Hipofaringe é um crescimento descontrolado de células malignas na parte de trás da faringe, formando um câncer, denominado carcinoma espinocelular. Os fatores de risco para o desenvolvimento desse distúrbio é o uso de tabaco, álcool e uma dieta desprovida de nutrientes (Cleveland Clinic, s.d.).

2.3.17. Tumor na Laringe

O desenvolvimento de células malignas na laringe é denominado Tumor na Laringe e pode ser dividido em três tipos: supra glótico, glótico e subglótico. Esse distúrbio torna as cordas vocais fixas, levando a um fechamento incompleto da glote, o que resulta numa rouquidão (Mayo Clinic, s.d.). A combinação entre tabaco e álcool é um dos principais fatores de risco para essa patologia.

2.4. Parâmetros dos Sinais Acústicos

Esta subdivisão aborda os principais parâmetros na análise acústica extraídos de gravações dispostas na base de dados SVD. Esses elementos são as perturbações em frequência e amplitude, proporções harmônico-ruído e autocorrelação.

Com o processamento de sinais é possível extrair um conjunto de parâmetros que colaboram para a detecção de distúrbios na laringe. Para isso, é necessário que haja um sinal gravado de forma constante tanto de indivíduos com uma voz considerada saudável quanto pacientes que possuem algum distúrbio na voz. Assim, é possível retirar informações relevantes, como os parâmetros acústicos. Nesta dissertação foram escolhidos os parâmetros *jitter* absoluto, *jitter* relativo, *shimmer* absoluto, *shimmer* relativo, HNR, NHR e autocorrelação por serem uma linha de pesquisa já conhecida e estudada durante o período de mestrado.

Para este estudo, esses atributos já passaram pelo processamento de sinais e são adquiridos por meio da base de dados curada implementada por Fernandes (2018), com as vogais sustentadas /a/, /i/ e /u/ em tons alto, baixo e normal.

2.4.1. Perturbações em Frequência

O ***jitter absoluto (jitta)***, como aponta Teixeira, Fernandes, Teixeira & Fernandes, (2018), “é a variação do período glotal entre os ciclos, ou seja, a média da diferença absoluta entre períodos consecutivos”. Esta descrição pode ser apresentada em forma de equação, como pode ser visto na Equação (2.1).

$$jitta = \frac{1}{N-1} \sum_{i=2}^{N-1} |T_i - T_{i-1}| \quad (2.1)$$

O ***jitter relativo (jitter)***, para Teixeira et al. (2018), “é a diferença absoluta média entre períodos glóticos consecutivos divididos pelo período médio e expressos como porcentagem”, como é apresentado na Equação (2.2).

$$jitter = \frac{\frac{1}{N-1} \sum_{i=2}^N |T_i - T_{i-1}|}{\frac{1}{N} \sum_{i=1}^N T_i} \times 100 \quad (2.2)$$

As variáveis dispostas nas expressões que designam a perturbação em frequência do sinal de fala são: o tamanho do período glotal, T_i , e o número total de períodos glotais, N .

2.4.2. Perturbações em amplitude

O ***shimmer absoluto (ShdB)*** refere-se à variação da amplitude pico-a-pico de uma onda sonora, em decibéis, extraída de um sinal gravado que gera uma vogal extensa e constante. Esse parâmetro, como Fernandes (2018) aponta, pode ser enunciado como a média absoluta da

multiplicação entre uma constante de valor 20 e o logaritmo de base 10 da razão entre 2 períodos consecutivos, como pode ser visto na Equação (2.3).

$$ShdB = \frac{1}{N-1} \sum_{i=1}^{N-1} \left| 20 \times \log \left(\frac{A_{i+1}}{A_i} \right) \right| \quad (2.3)$$

O *shimmer relativo (Shim)*, expressa pela Equação (2.4) é um dos parâmetros para análise de doenças que afetam a qualidade vocal dos pacientes e é determinado, de acordo com (Teixeira et al., 2018) “como a média absoluta da diferença entre as amplitudes de períodos consecutivos divididos pela amplitude média, expressa em porcentagem”.

$$Shim = \frac{\frac{1}{N-1} \sum_{i=2}^N |A_{i+1} - A_i|}{\frac{1}{N} \sum_{i=1}^N A_i} \times 100 \quad (2.4)$$

Nas equações utilizadas para definir as perturbações em amplitude foram utilizadas as variáveis A_i , que designa a magnitude do período glotal e N , que é o número total de períodos glotais.

2.4.3. Autocorrelação

A correlação de 2 sinais, por si só, é a soma dos valores de seus produtos. Logo, a autocorrelação pode ser elucidada como a correlação de um sinal consigo mesmo. A função da autocorrelação, para uma onda sonora, é um método de detectar a periodicidade do sinal, em que se divide a autocorrelação da janela normalizada do sinal selecionado pela autocorrelação normalizada da janela utilizada, como pode ser visto na Equação (2.5). De acordo com Teixeira et al. (2018), essa propriedade “fornece uma medida das partes de fala semelhantes repetidas ao longo do sinal”.

$$r_x(\tau) = \frac{r_a(\tau)}{r_w(\tau)} \quad (2.5)$$

Na Equação (2.5) a expressão $r_a(\tau)$ é a autocorrelação normalizada de parte do sinal selecionado e $r_w(\tau)$ é a autocorrelação normalizada da janela de *Hanning*. Como é apontado por Fernandes (2018), a autocorrelação normalizada de parte do sinal é fornecida pela Equação (2.6), enquanto a autocorrelação normalizada da janela de *Hanning* é expressa pela Equação (2.7).

$$r_a(\tau) = r_a(-\tau) \frac{\int_0^{T-\tau} a(t)a(t+\tau)dt}{\int_0^T a^2(t)dt} \quad (2.6)$$

$$r_w(\tau) = \left(1 - \frac{|\tau|}{T}\right) \left(\frac{2}{3} + \frac{1}{3} \cos \frac{2\pi\tau}{T}\right) + \frac{1}{2\pi} \sin \frac{2\pi|\tau|}{T} \quad (2.7)$$

As variáveis das equações anteriormente citadas são: T , que é a duração do sinal, $a(t)$, que compreende à uma janela do sinal, e τ , que é uma função simétrica ao atraso.

2.4.4. HNR

Ao pesquisar sobre as características harmônicas da voz, deparamo-nos com o parâmetro *Harmonic to Noise Ratio*, que trabalha com a relação entre os componentes periódicos e aperiódicos de um segmento de fala, sendo o primeiro componente decorrente da vibração das cordas vocais, enquanto o segundo componente é um ruído glótico. Vários autores indagam diversas formas de expressar o HNR. Neste trabalho será seguida a sugestão apresentada por Teixeira et al. (2018), explícita na Equação (2.8), que “consiste em medir a energia do primeiro pico da autocorrelação normalizada e considerar que essa é a energia da componente harmônica do sinal e considerar a energia remanescente como a energia do ruído dada pela diferença entre 1 e a energia harmônica”.

$$HNR(dB) = 10 \times \log_{10} \frac{r'_x(\tau_{max})}{1 - r'_x(\tau_{max})} \quad (2.8)$$

Nesta equação o termo $r'_x(\tau_{max})$ é o máximo local da autocorrelação normalizada.

2.4.5. NHR

Antagonicamente ao HNR, *Noise to Harmonic Ratio* é a relação entre a componente aperiódica, mais especificamente o ruído, e a componente periódica, relacionada a vibração das cordas vocais (Teixeira et al., 2018), (Fernandes, 2018). Como esse parâmetro não é medido no domínio logarítmico seus valores seguem direções opostas e não são propriamente inversos ao HNR. A Equação (2.9), que descreve o NHR, é determinada em função da autocorrelação.

$$NHR = 1 - \text{autocorrelação} \quad (2.9)$$

Capítulo 3: Análise Estatística Descritiva

Neste capítulo será discutido um método de análise estatístico descritivo para averiguar se os grupos extraídos da base de dados SVD podem se agrupar.

3.1. Contextualização

Trabalhar com redes neurais não-supervisionadas, como é o caso desse trabalho, implica-se na necessidade de maiores informações a respeito dos elementos da base de dados a ser mapeada. A obtenção dessas informações pode ser feita por meio da análise estatística ou da opinião de um especialista. Para esta dissertação o recurso aplicado a fim de alcançar uma suposição inicial é a análise estatística descritiva. Mais especificamente o Diagrama de Caixa, por ser um método já investigado em outros artigos (Oliveira et al, 2020), (Teixeira et al, 2018).

A estatística nada mais é do que a transformação de informações reais em fatos numéricos. Os métodos estatísticos nos auxiliam na tomada de decisão inteligente através do estudo numérico. Esses métodos podem se dividir em: descritivo e inferencial. Na estatística descritiva as informações da base de dados são geralmente grandes, sendo crucial a utilização de técnicas de organização e exibição para entender o conteúdo examinado. Na estatística inferencial, os métodos fornecem informações para ajudar na tomada de decisões (Mann, 2010).

O Diagrama de Caixa é um método de análise estatística descritiva que, basicamente, compara os dados através de 3 situações que as caixas podem estar. Neste trabalho, utiliza-se a ferramenta de Diagrama de Caixa presente no *software* MATLAB®.

3.2. Diagrama de Caixa

O Diagrama de Caixa é uma representação gráfica dos elementos, que proporciona uma ideia da posição, dispersão, assimetria e pontos divergentes de um conjunto de dados (Mann, 2010), (Morettin & Bussab, 2010). Para iniciar a construção desse diagrama, deve-se classificar os dados em ordem crescente e calcular os valores dos quartis e do intervalo interquartil. O

termo quartil é utilizado para expressar 3 medidas quando se divide um conjunto de dados em quatro partes iguais (Mann, 2010), (Morettin & Bussab, 2010). O primeiro quartil (Q_1) é o valor que engloba 25% das observações do conjunto, o segundo quartil (Q_2) é a mediana e contém 50% dos elementos, já o terceiro quartil (Q_3) compreende 75% dos dados. O intervalo interquartil (IQR), expresso pela Equação (3.1), é a diferença entre o terceiro e primeiro quartis.

$$IQR = Q_3 - Q_1 \quad (3.1)$$

Para prosseguir, deve-se encontrar os limites superior e inferior do diagrama, que são dados pelos pontos $LS = Q_3 + (1,5 \times IQR)$ e $LI = Q_1 - (1,5 \times IQR)$, respectivamente. Assim, pode-se desenhar um retângulo, que se inicia no primeiro quartil e termina no terceiro quartil, com uma reta onde a mediana está localizada. De acordo com Mann (2010), uma linha horizontal é feita no menor e no maior valor dentro dos limites estabelecidos anteriormente, designando as cercas internas. Assim, faz-se uma linha ligando essas cercas ao retângulo, obtendo o diagrama completo. A Figura 2 mostra a representação gráfica de um Diagrama de Caixa.

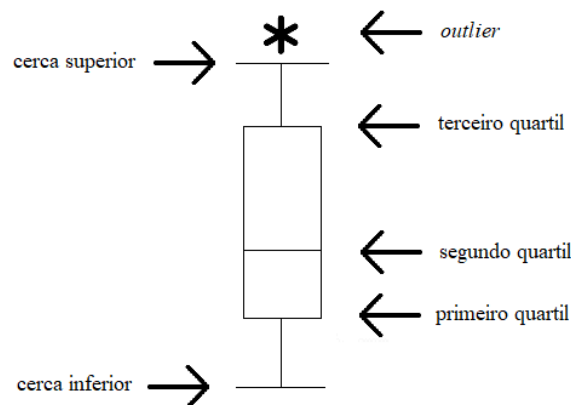


Figura 2 – Representação gráfica da construção de um Diagrama de Caixa.

Qualquer valor estabelecido para fora das cercas definidas está discrepante com o resto dos dados e deve ser considerado um *outlier* (Mann, 2010), (Morettin & Bussab, 2010).

3.3. Conferência entre Pares de Caixas

Sabendo que o Diagrama de Caixa é usado para uma análise estatística descritiva, essa ferramenta, como aponta Teixeira et al. (2018), essencialmente compara os elementos A e B através de 3 situações. Na primeira situação, exposta na Figura 3, não há sobreposição entre as caixas, sendo B maior que A. Assim, há uma diferença entre os grupos A e B. A Figura 4 mostra a segunda situação, em que as caixas se sobrepõem sem que ambas medianas sobreponham as caixas, logo, provavelmente, há uma diferença entre os grupos A e B. Na terceira situação, apresentada pela Figura 5, as caixas se sobrepõem e suas medianas sobrepõem as caixas, determinando que nenhuma diferença pode ser atribuída aos grupos A e B.

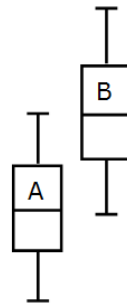


Figura 3 – Primeira situação: há uma diferença entre os grupos. Fonte: Teixeira et al. (2018).

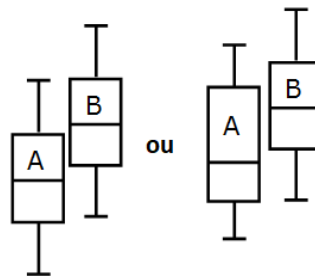


Figura 4 – Segunda situação: há, provavelmente, uma diferença entre os grupos. Fonte: Teixeira et al. (2018).

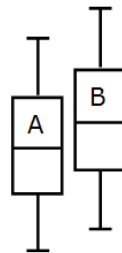


Figura 5 – Terceira situação: não há diferença entre os grupos. Fonte: Teixeira et al. (2018).

3.4. Metodologia

Cada conjunto investigado nessa dissertação possui amostras com tamanhos variados, como é possível visualizar na Tabela 1, sendo 830 indivíduos que somam um total de 7470 amostras, pois, cada sujeito possui um ficheiro pronunciando as vogais /a/, /i/ e /u/ sustentadas nos tons alto, baixo e normal.

Por meio da base de dados curada implementada por Fernandes (2018), é possível obter os parâmetros *jitta*, *jitter*, ShdB, Shim, HNR, NHR e autocorrelação. Essa base de dados se encontra estruturada em uma tabela, em que as colunas indicam os atributos e as linhas são designadas pelas amostras em si. O grupo de Controlo compõe os primeiros elementos desse quadro, sendo ordenados em sexo feminino seguido do sexo masculino. Em sequência, repete-se a mesma estrutura para todas as patologias em ordem alfabética: Carcinoma, Cisto, Disfonia Espasmódica, Disfonia Funcional, Disfonia Hiperfuncional, Disfonia Hipofuncional, Disfonia Hipotônica, Disfonia Psicogênica, Edema de Reinke, Granuloma, Granuloma por Intubação, Laringite Crônica, Paralisia das Cordas Vocais, Pólipos das Cordas Vocais, Tumor na Hipofaringe e Tumor na Laringe.

Por meio da ferramenta Diagrama de Caixa, disposta no *software* MATLAB®, efetua-se uma análise estatística descritiva comparando cada grupo de teste mediante aos parâmetros abordados anteriormente. Com o intuito de realizar uma investigação completa entre todos os conjuntos, faz-se uma comparação entre pares desses conjuntos para todas as propriedades, a fim de notar uma das 3 situações supracitadas. Assim, torna-se possível descobrir se os conjuntos se enquadram na terceira situação, para agrupá-los em um único conjunto, ou se eles formam agrupamentos distintos.

3.5. Resultados e Discussões

A análise estatística descritiva, comparando cada grupo de teste, se dividiu em 8 componentes, a partir da quantidade de parâmetros acústicos. Visto que a comparação para o *jitter* absoluto foi feita separando os gêneros, pois, de forma geral, vozes masculinas têm baixa frequência fundamental e, conseqüentemente, períodos glotais maiores. Logo, para períodos glotais mais longos, admitem-se desvios maiores (Teixeira et al., 2018).

Nesta investigação, para cada parâmetro, uma tabela de cores é utilizada para uma fácil visualização a respeito do possível agrupamento das amostras. Ao comparar todos os grupos

entre si dispostos no banco de dados, pode ser feita uma análise através do Diagrama de Caixa, enquadrando as amostras nas 3 situações previamente discutidas. A cor vermelha indica a primeira situação, em que não é possível agrupar os elementos; a cor amarela implica na segunda situação, em que os conjuntos, possivelmente, não se aglomeram; e a cor verde aponta a terceira categoria, onde os conjuntos se agrupam.

3.5.1. *Jitter* Absoluto Feminino

A Figura 6 mostra a comparação do Diagrama de Caixas para o *jitter* absoluto de indivíduos do gênero feminino. Como é possível perceber pela figura que não há dados suficientes a respeito do Tumor na Hipofaringe, da Disfonia Hipotônica e do Granuloma por Intubação, portanto, esses elementos não são analisados nesse atributo.

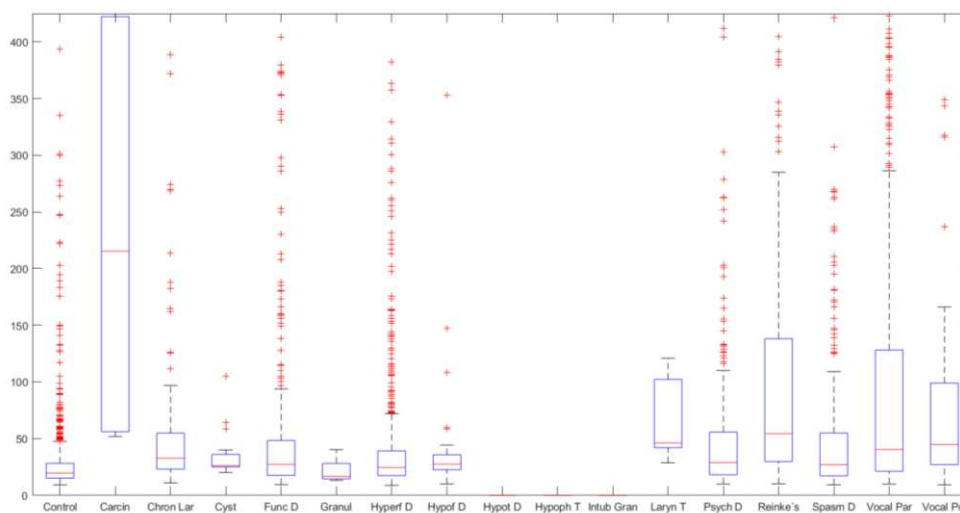
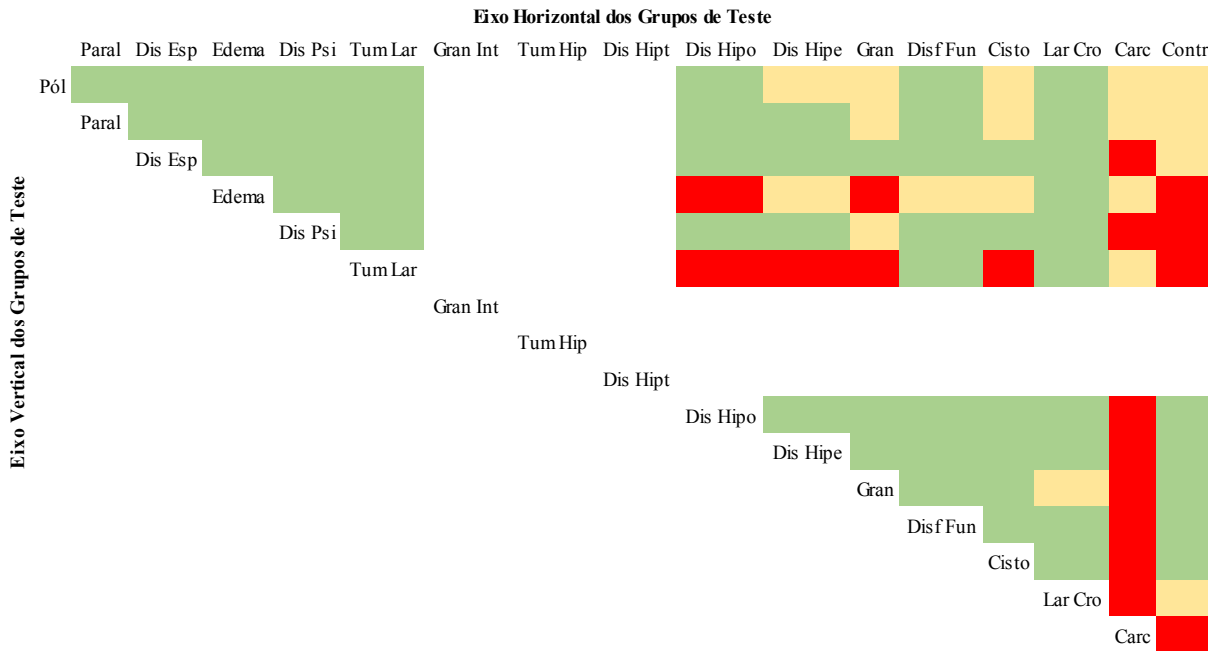


Figura 6 – Diagrama de Caixa comparando os conjuntos da base de dados por meio do *jitter* absoluto para o gênero feminino.

A partir da Figura 6 pode-se comparar os grupos, tendo como base as informações a respeito das comparações entre caixas dispostas no início desse capítulo, e criar a Tabela 2, com os resultados dessa investigação.

Tabela 2 – Resultado das semelhanças entre os conjuntos com base no jitter absoluto feminino.



De acordo com a Tabela 2, as disfunções Pólipo, Paralisia nas Cordas Vocais, Disfonia Espasmódica, Edema de Reinke, Disfonia Psicogênica e Tumor na Laringe podem se agrupar num grande grupo. A Disfonia Hipofuncional, Disfonia Hiperfuncional, Disfonia Funcional, Cisto, Laringite Crónica e o grupo de Controlo se unem num outro conjunto. O Carcinoma, neste parâmetro, não se agrupa a nenhum conjunto, enquanto a Disfonia Espasmódica, Disfonia Funcional, Disfonia Hiperfuncional e Laringite Crónica se juntam a maioria dos grupos.

3.5.2. Jitter Absoluto Masculino

Com a Figura 7, que mostra o parâmetro *jitter* absoluto para o gênero masculino, criou-se a Tabela 3, que faz a comparação de todos os conjuntos utilizados do banco de dados.

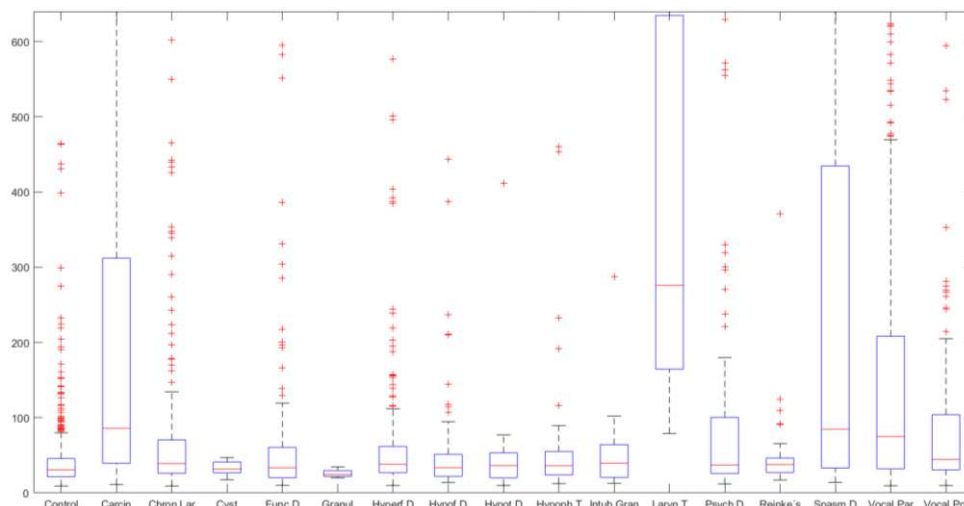
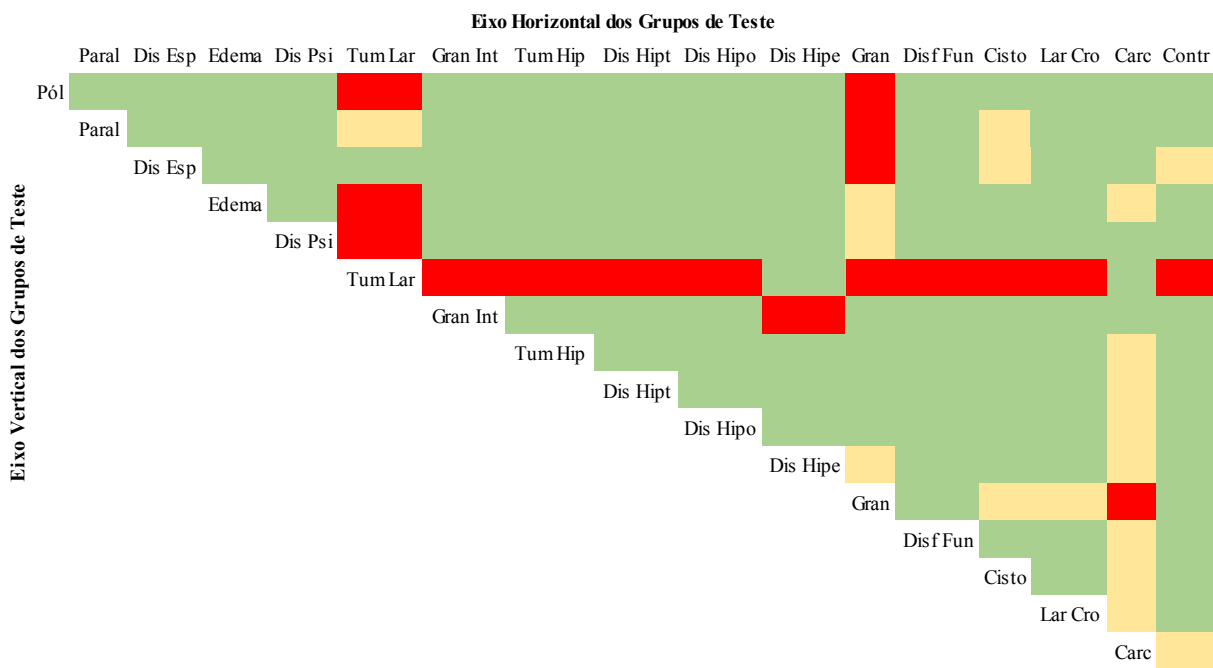


Figura 7 – Diagrama de Caixa comparando os conjuntos da base de dados por meio do *jitter* absoluto para o gênero masculino.

Tabela 3 – Semelhança entre os conjuntos com base no Diagrama de Caixa.



Na Tabela 3 é possível visualizar que o Tumor na Laringe não se agrupa a quase nenhum conjunto, com exceção da Disfonia Hiperfuncional, Carcinoma e Disfonia Espasmódica. Ao passo que a maioria das amostras são capazes de se agrupar entre si. O Pólipo nas Pregas Vocais se agrupa a todos os conjuntos, exceto ao Granuloma e Tumor na Laringe. O Granuloma por Intubação se une a todos os grupos, com exceção do Tumor na Laringe e Disfonia Hipotônica.

Um agrupamento pode ser feito pelo Pólipo, Paralisia, Disfonia Espasmódica, Edema de Reinke, Disfonia Psicogênica, Granuloma por Intubação, Tumor na Hipofaringe, Disfonia Hipotônica, Disfonia Hipofuncional, Disfonia Hiperfuncional, Disfonia Funcional e Laringite Crônica. O grupo de Controle, a Laringite Crônica, o Cisto, a Disfonia Funcional, a Disfonia Hiperfuncional, a Disfonia Hipofuncional, a Disfonia Hipotônica e o Granuloma por Intubação também se unem.

3.5.3. *Jitter* Relativo

A Figura 8 apresenta a investigação do parâmetro *jitter* relativo por meio do Diagrama de Caixa com todas as amostras dos grupos de teste.

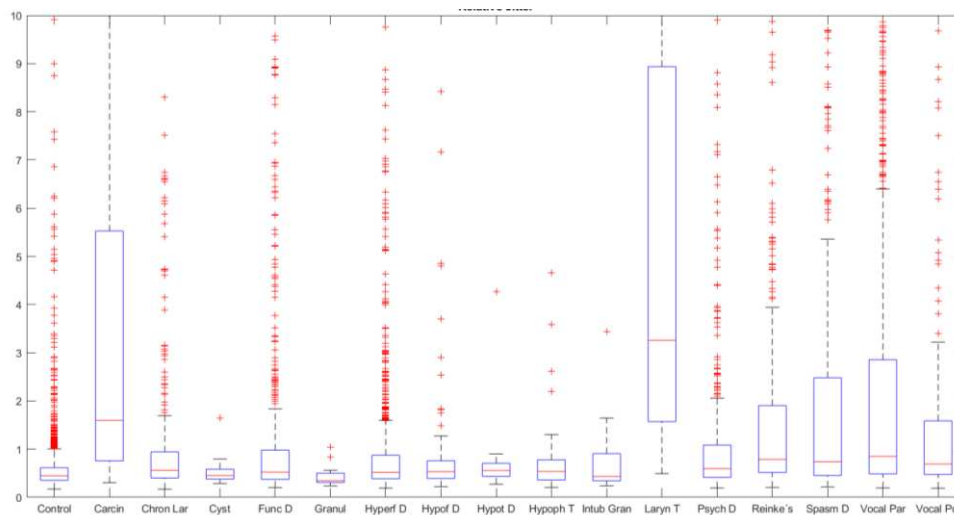
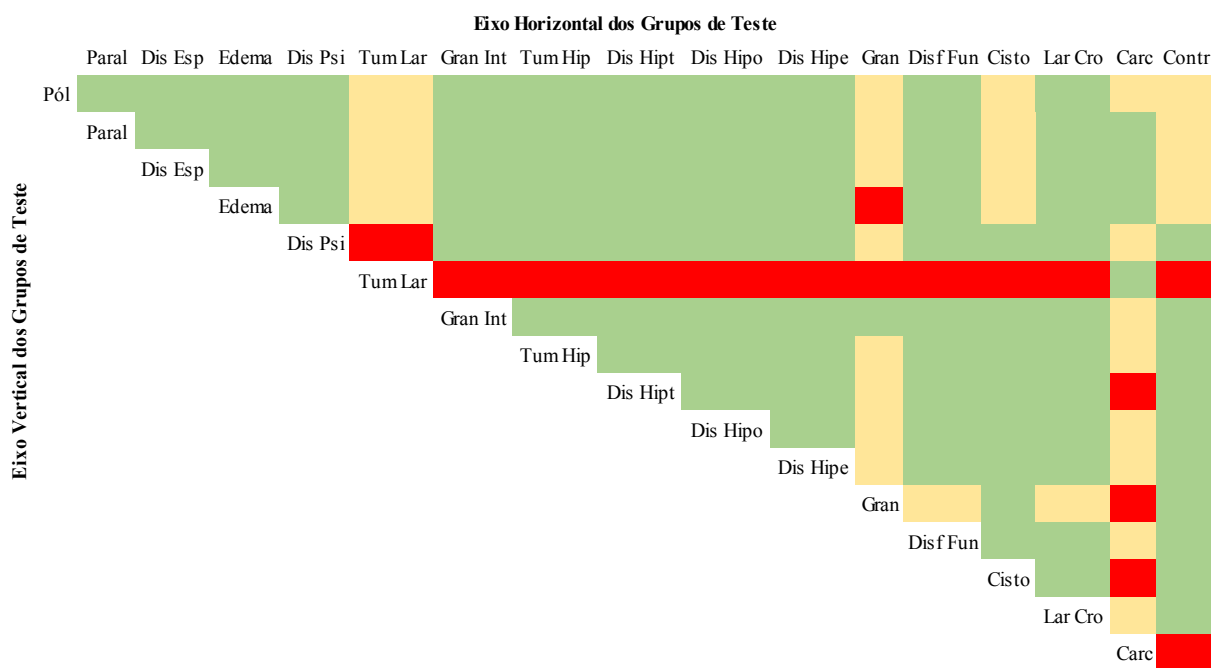


Figura 8 – Diagrama de Caixa do *jitter* relativo feito entre os conjuntos de teste.

Através da Figura 8 e do conhecimento a respeito das comparações entre as caixas explicitado neste capítulo, é possível apresentar, de forma sucinta, a comparação entre os grupos estudados neste parâmetro e suas classificações por meio da Tabela 4.

Tabela 4 – Resultados das comparações por meio do Diagrama de Caixa para o jitter relativo.



De maneira semelhante ao *jitter* absoluto do gênero masculino, os conjuntos Pólipo, Paralisia, Disfonia Espasmódica, Edema de Reinke, Disfonia Psicogênica, Granuloma por Intubação, Tumor na Hipofaringe, Disfonia Hipotônica, Disfonia Hipofuncional, Disfonia Hiperfuncional, Disfonia Funcional e Laringite Crônica se agrupam num único grupo. Outro agrupamento pode ocorrer pela união do grupo de Controle, Laringite Crônica, Disfonia Funcional, Disfonia Hiperfuncional, Disfonia Hipofuncional, Disfonia Hipotônica, Granuloma por Intubação e Disfonia Psicogênica.

3.5.4. *Shimmer* Absoluto

O Diagrama de Caixa para o *shimmer* absoluto é visto na Figura 9, em que todos os conjuntos são plotados e comparados entre si.

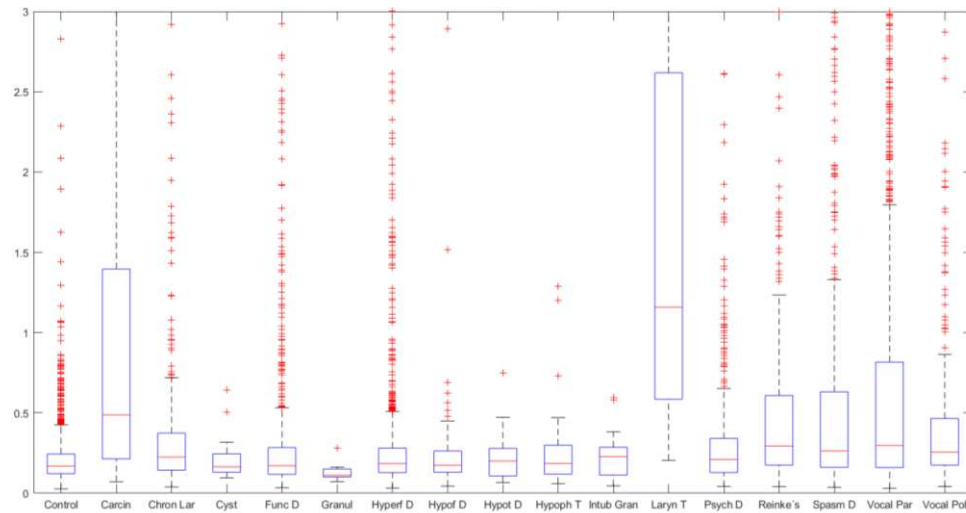
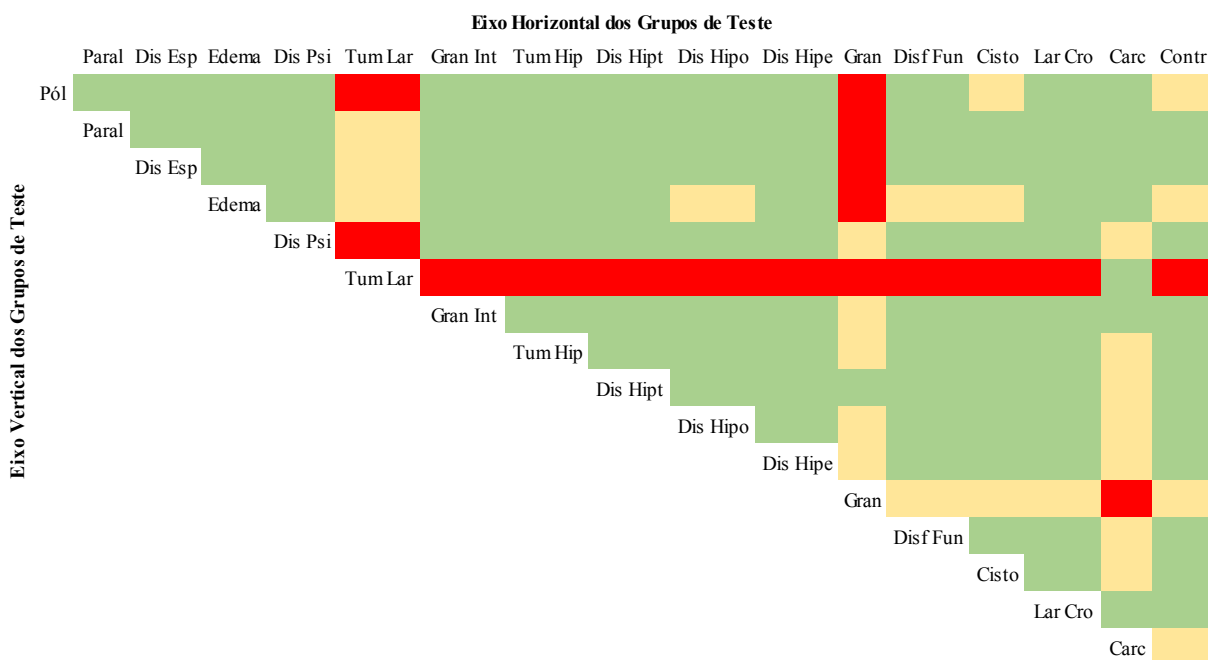


Figura 9 – Diagrama de Caixa comparando os conjuntos extraídos do SVD por meio do shimmer absoluto.

A comparação entre cada elemento da figura resulta numa situação, seja de agrupamento ou não. A Tabela 5 mostra o resultado dessas comparações, em que cada cor representa uma situação.

Tabela 5 – Resultado da comparação entre as caixas do parâmetro shimmer absoluto.



Por meio da tabela, é visto que os grupos Granuloma e Tumor na Laringe não se agrupam a quase nenhuma patologia. O Granuloma por Intubação, o Tumor na Hipofaringe, a Disfonia Hipotônica, a Disfonia Hipofuncional, a Disfonia Hiperfuncional, a Disfonia Funcional, o Pólipo, a Paralisia, a Disfonia Espasmódica e a Disfonia Psicogênica podem se

aglomerar num grande grupo. Outro possível grupo é formado pelo Granuloma por Intubação, Tumor na Hipofaringe, Disfonia Hipotônica, Disfonia Hipofuncional, Disfonia Hiperfuncional, Disfonia Funcional, Cisto, Laringite Crônica e Controlo.

3.5.5. *Shimmer* Relativo

O Diagrama de Caixa do *shimmer* relativo pode ser visualizado na Figura 10, por onde constrói-se a Tabela 6, que informa, de maneira sucinta, as aglomerações feitas por meio desse parâmetro.

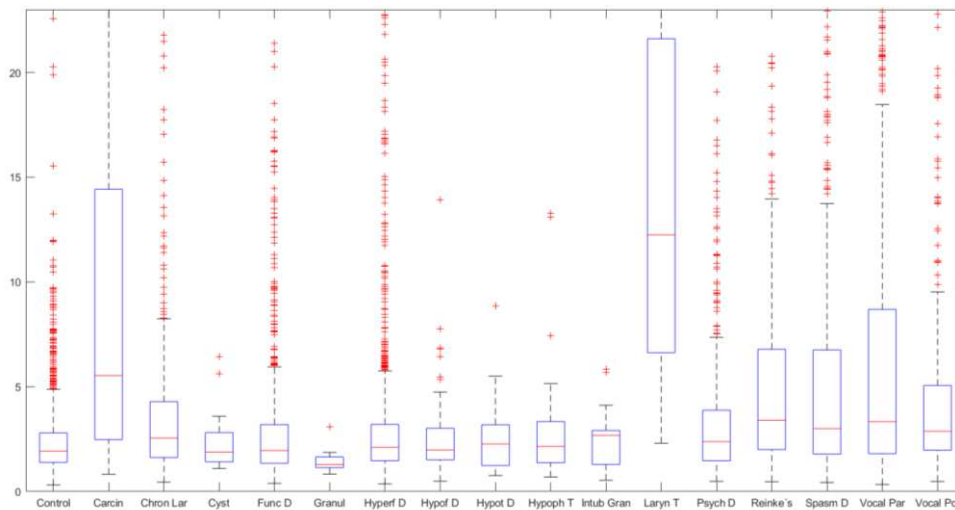
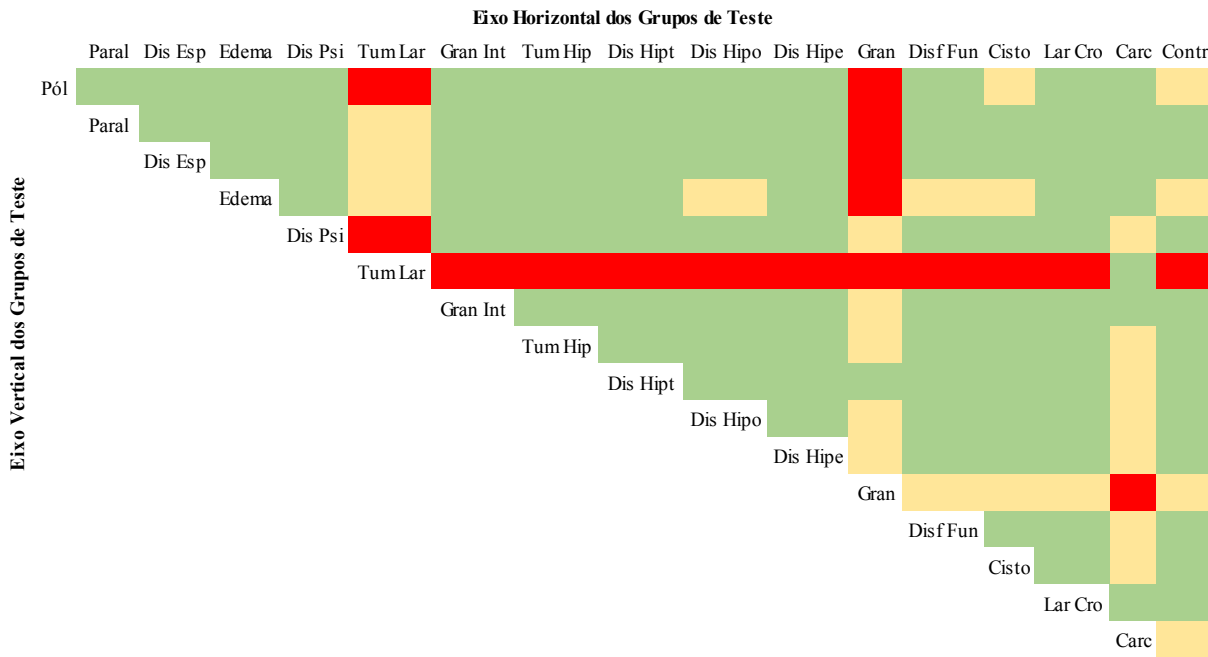


Figura 10 – Diagrama de Caixa comparando os conjuntos por meio do *shimmer* relativo.

Tabela 6 – Possíveis agrupamentos por meio do Diagrama de Caixa do shimmer relativo.



A Tabela 6 é exatamente igual ao quadro do parâmetro anterior, implicando no agrupamento do Granuloma por Intubação, o Tumor na Hipofaringe, a Disfonia Hipotônica, a Disfonia Hipofuncional, a Disfonia Hiperfuncional, a Disfonia Funcional, o Pólipo, a Paralisia, a Disfonia Espasmódica e a Disfonia Psicogênica e na união do Granuloma por Intubação, Tumor na Hipofaringe, Disfonia Hipotônica, Disfonia Hipofuncional, Disfonia Hiperfuncional, Disfonia Funcional, Cisto, Laringite Crônica e Controle. Os conjuntos Granuloma e Tumor na Laringe não se agrupam a nenhum grupo, com exceção da Disfonia Hipofuncional para o Granuloma e do Carcinoma para o Tumor na Laringe.

3.5.6. HNR

O Diagrama de Caixa do *Harmonic to Noise Ratio* pode ser visto na Figura 11. Essa figura serve como base para a construção da Tabela 7, que apresenta os valores das comparações entre cada grupo do diagrama, fornecendo uma rápida visualização a respeito do agrupamento das amostras.

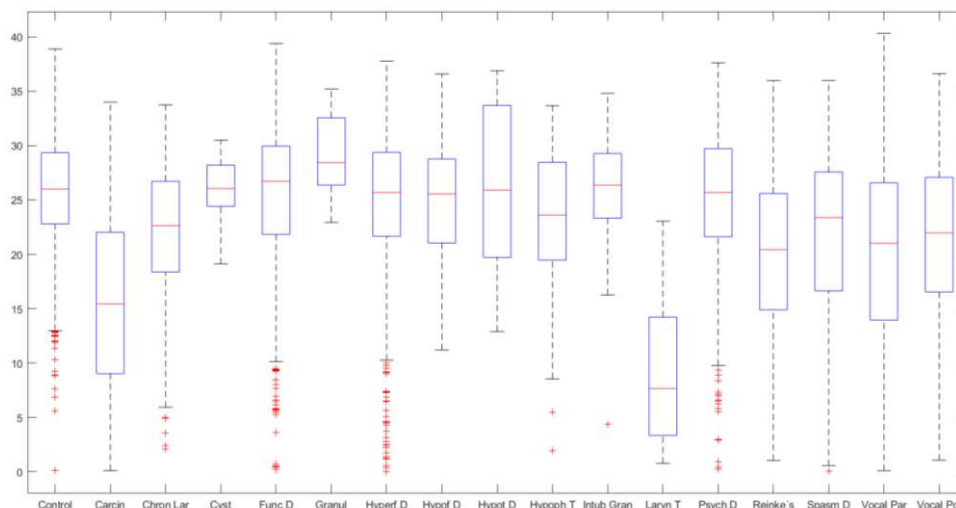
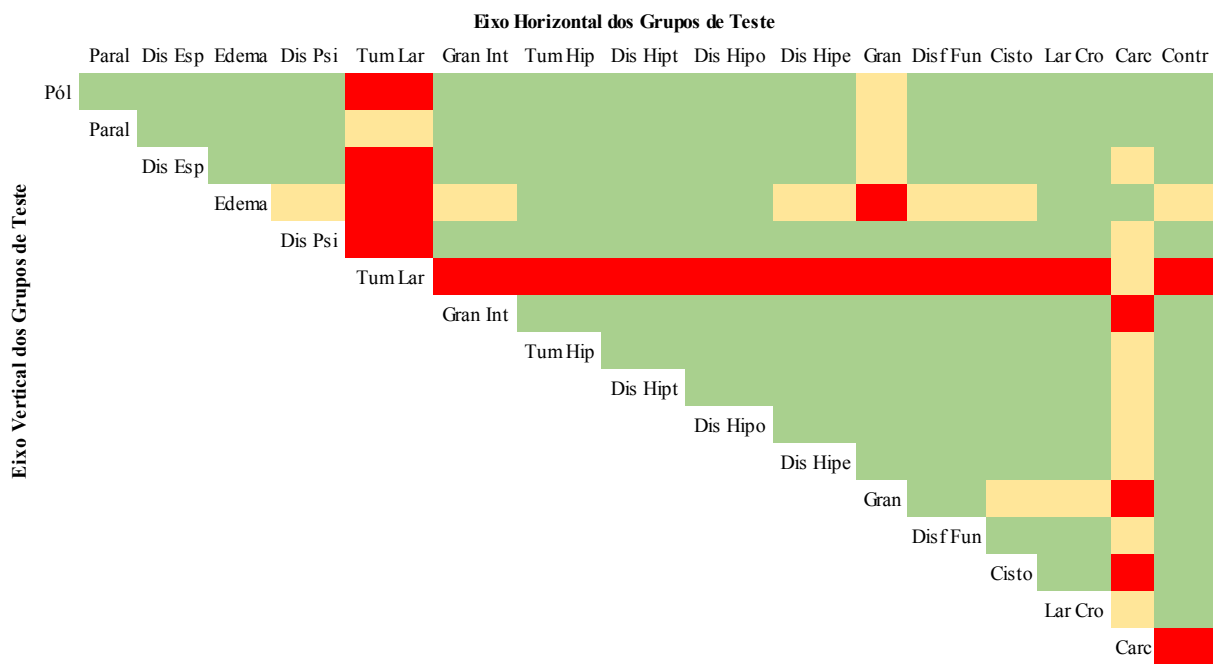


Figura 11 – Diagrama de Caixa utilizando o HNR para analisar os conjuntos do banco de dados.

Tabela 7 – Resultado do Diagrama de Caixa utilizando as amostras do HNR sobre semelhanças entre os grupos.



Com o suporte da Tabela 7, pode-se ver que o Tumor na Laringe não se agrupa a nenhum conjunto. Em compensação, é possível visualizar um grande grupo composto pelo Pólipo, Paralisia, Disfonia Espasmódica, Disfonia Psicogênica, Granuloma por Intubação, Tumor na Hipofaringe, Disfonia Hipotônica, Disfonia Hipofuncional, Disfonia Hiperfuncional, Disfonia Funcional, Cisto, Laringite Crônica e Controlo.

3.5.7. NHR

A Figura 12 mostra os 17 grupos de teste para o parâmetro NHR. Com essa figura, faz-se a comparação entre cada conjunto com base na teoria do Diagrama de Caixa, construindo a Tabela 8.

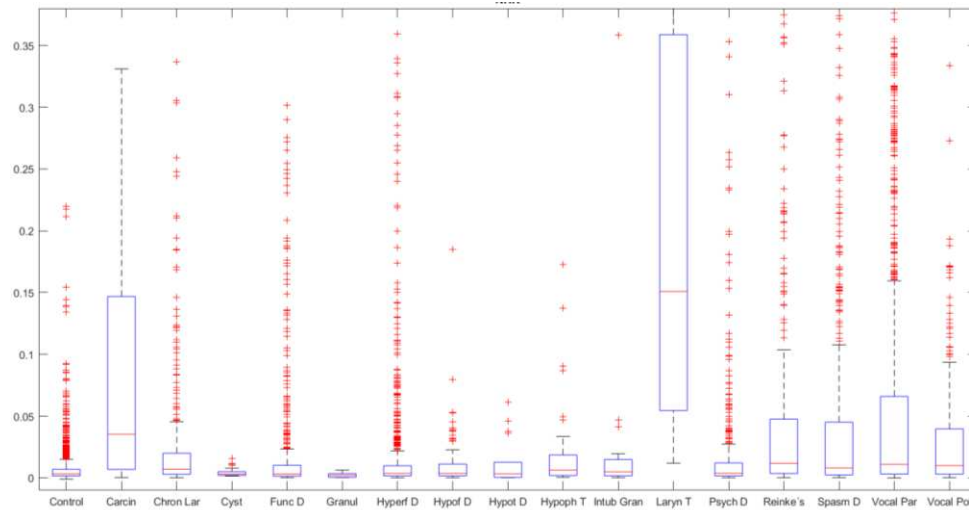
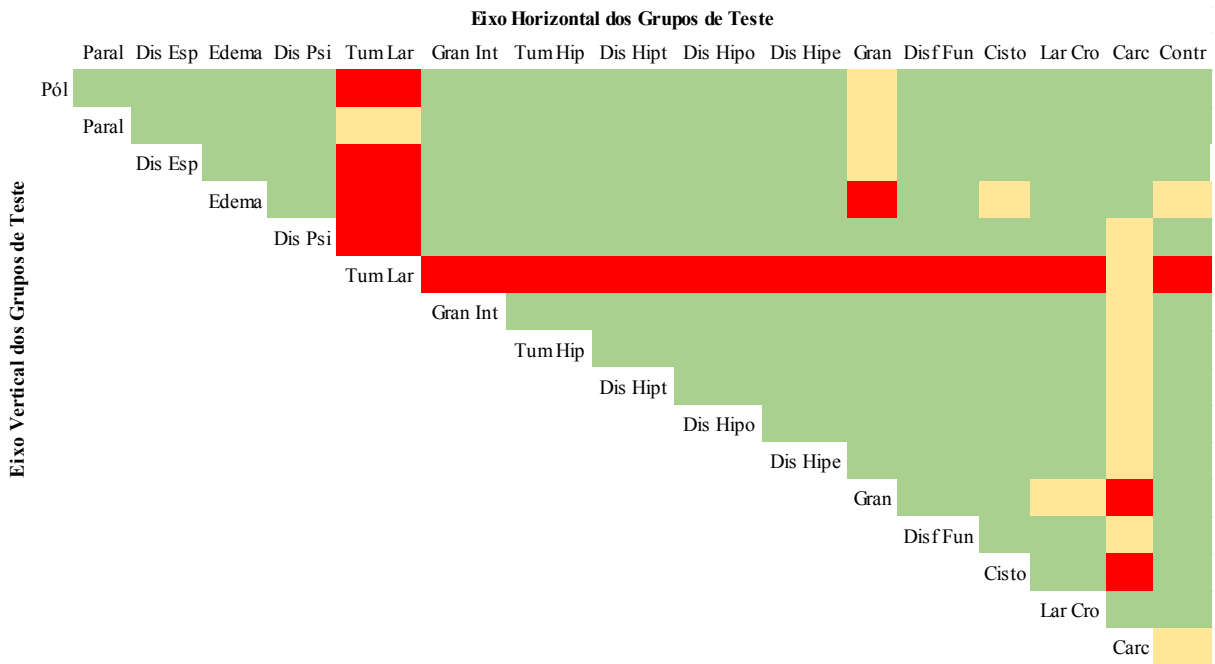


Figura 12 – Comparação dos conjuntos através do Diagrama de Caixa utilizando o NHR.

Tabela 8 – Resultados das comparações por Diagrama de Caixa dos conjuntos de teste através do parâmetro NHR.



Por meio da Tabela 8, nota-se que, assim como no HNR, um grande conjunto pode ser criado com o Pólipo, Paralisia, Disfonia Espasmódica, Disfonia Psicogênica, Granuloma por Intubação, Tumor na Hipofaringe, Disfonia Hipotônica, Disfonia Hipofuncional, Disfonia Hiperfuncional, Disfonia Funcional, Cisto, Laringite Crônica e Controle. Neste parâmetro, o Tumor na Laringe também não se agrupa a nenhum conjunto.

3.5.8. Autocorrelação

No atributo de autocorrelação, disposto na Figura 13, compara-se todas as caixas no diagrama para verificar a relação entre cada um dos 17 conjuntos. Essa comparação pode ser vista na Tabela 9, que mostra o resultado de cada comparação com base no item 3.3.

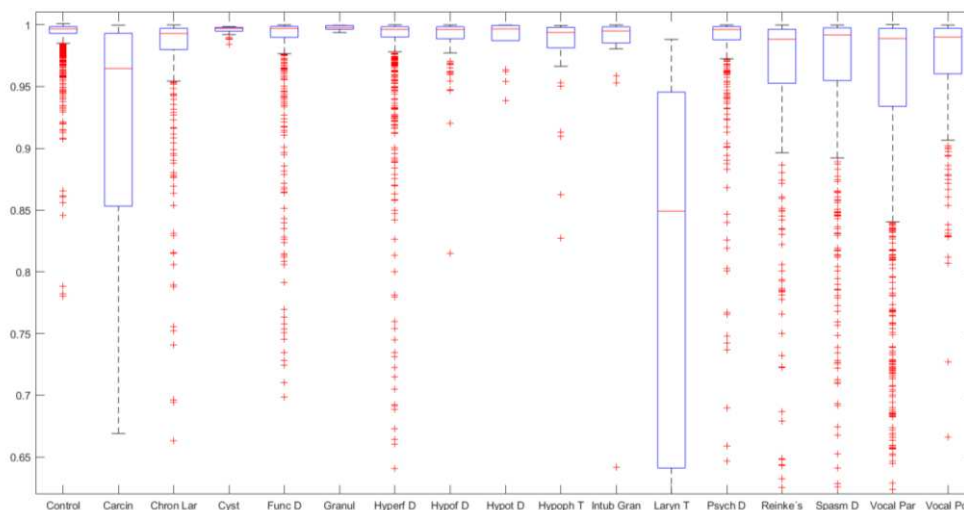
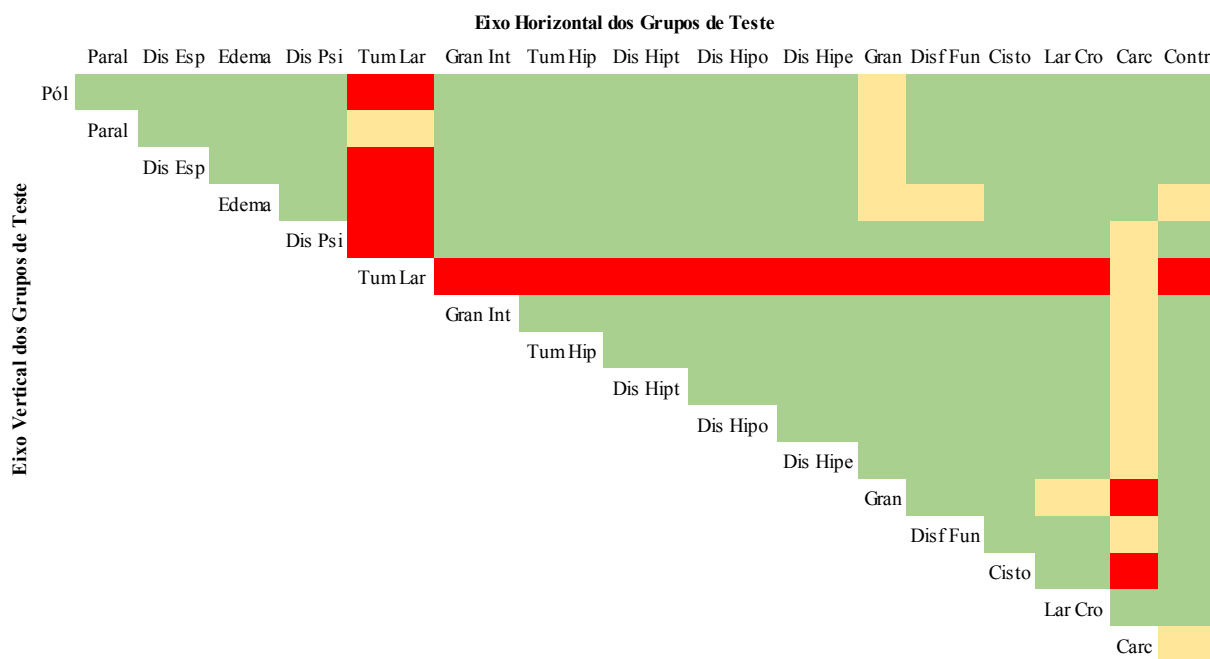


Figura 13 - Utilização da autocorrelação no estudo do Diagrama de Caixa para os grupos de teste.

Tabela 9 – Resultado das semelhanças entre os grupos de teste por meio da autocorrelação.

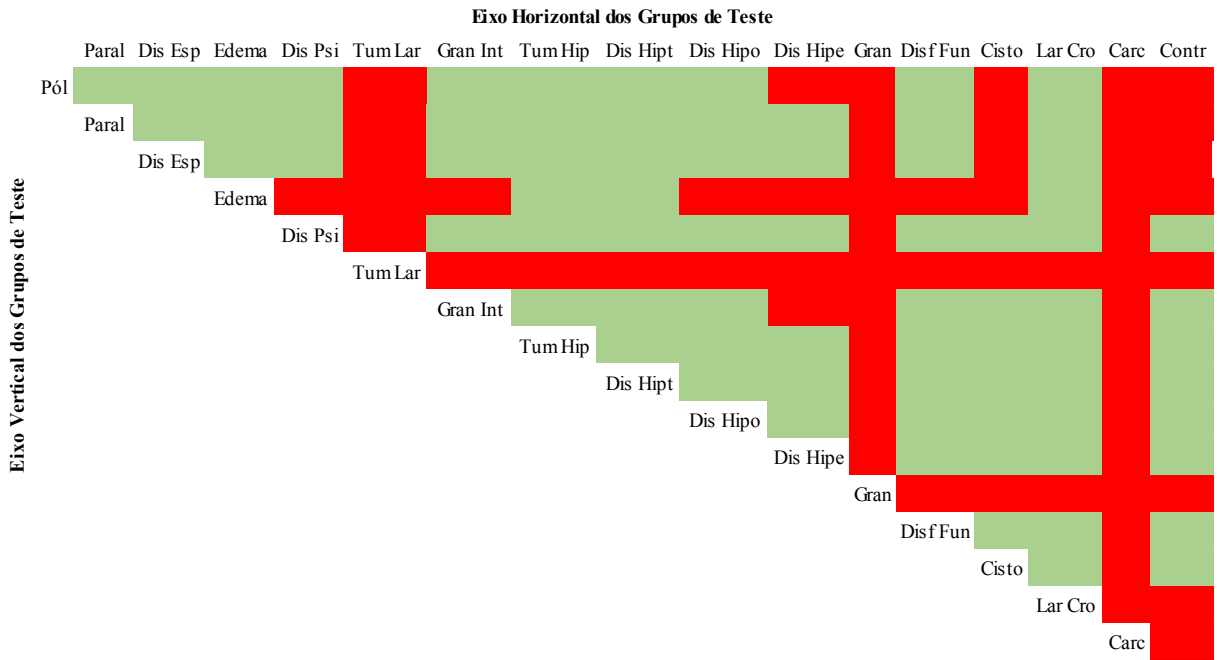


Por meio dos resultados expressos na Tabela 9, percebe-se que o Tumor na Laringe não se agrupa a nenhum conjunto. Ao passo que o Pólipo, Paralisia, Disfonia Espasmódica, a Disfonia Psicogênica, Granuloma por Intubação, Tumor na Hipofaringe, Disfonia Hipotônica, Disfonia Hipofuncional, Disfonia Hiperfuncional, Cisto e Laringite Crônica se juntam num conjunto. Outro conjunto pode ser formado pelo grupo de Controle, Laringite Crônica, Cisto, Disfonia Funcional, Disfonia Hiperfuncional, Disfonia Hipofuncional, Disfonia Hipotônica, Tumor na Hipofaringe e Granuloma por Intubação.

Ao investigar cada propriedade para cada par de conjuntos por meio das comparações, é possível organizar a Tabela 10, que apresenta como estrutura um triângulo. Nela estão descritas as relações entre os grupos de teste, no qual cada parâmetro foi analisado para uma parêntese dos conjuntos. Assim, se para todos os fatores, essa parêntese apresentou os mesmos resultados ou um mesmo valor para a maior parcela dessas comparações, assume-se essa medida como valor final. A menos que seja a terceira situação, que precisa estar presente em todos os atributos para considerar os conjuntos agrupáveis.

Para melhor entendimento, os valores da tabela estão adaptados a 2 possíveis situações. A cor verde indica que os conjuntos podem se agrupar, enquanto a cor vermelha implica que os conjuntos não podem se unir devido a divergências.

Tabela 10 - Relações entre os grupos de teste com base na comparação de Diagrama de Caixas.



3.6. Conclusão

Com averiguação de similaridades entre os grupos dispostos na base de dados implementada por Fernandes (2018) através da análise estatística descritiva é possível perceber que várias análises podem ser feitas de maneiras diferentes, obtendo, inclusive, *clusters* diferentes.

Como é visto na subseção 3.5.1, no *jitter* absoluto feminino, a Disfonia Hipotônica, o Tumor na Hipofaringe e o Granuloma por Intubação não apresentam dados. Com isso, ao fazer a comparação entre as patologias são considerados apenas 6 parâmetros para esse caso.

Ao observar a Tabela 10 constatou-se que, assim como é apresentado Oliveira, Dajer, Fernandes & Teixeira (2020), o Tumor na Laringe, o Carcinoma e o Granuloma não se agrupam com nenhum conjunto. No entanto, um conjunto pode ser formado com as patologias Disfonia Espasmódica, Paralisia, Pólipo, Disfonia Psicogênica, Granuloma por Intubação, Tumor na Hipofaringe, Disfonia Hipotônica, Disfonia Hipofuncional, Disfonia Funcional e Laringite Crônica. Outro grupo ocorre com a união do Controle, Laringite Crônica, Cisto, Disfonia Funcional, Disfonia Psicogênica, Tumor na Hipofaringe, Disfonia Hipotônica, Disfonia

Hipofuncional e Disfonia Hiperfuncional. Portanto, não levando em consideração os 3 grupos que não se unem a nenhum outro conjunto, há a existência de 2 possíveis agrupamentos, onde ocorre a intersecção de disfunções entre e esses grupos.

Capítulo 4: Métodos de *Clustering*

Nesta seção serão estudados os conceitos clássicos de agrupamento, que se dividem em 2 categorias: partição e hierarquia. Mais especificamente serão abordados os Mapas Auto-Organizáveis de Kohonen e processo hierárquico aglomerativo, respectivamente.

4.1. Contextualização

O estudo de agrupamentos é útil em diversas áreas da exploração de dados e, por conseguinte, vem crescendo rapidamente (Jain & Dubes, 1988), (Jain, Murty & Flynn, 1999), podendo ser aplicado em diversos campos da ciência. Para Jain e Dubes (1988), as técnicas computacionais fundamentadas nesse estudo apresentam critérios mais consolidados para formar aglomerações do que processos manuais.

A análise de agrupamentos é um método de organização de dados, separando-os em grupos com base em suas (dis)similaridades. Em função disso, objetos pertencentes a um mesmo aglomerado são muito semelhantes entre si, enquanto divergem de objetos de outros conjuntos (Gan et al., 2007), (Jain & Dubes, 1988), (Jain et al., 1999), (Berkhin, 2006). O objetivo dessa análise, conforme Jain e Dubes (1988) constatam, é encontrar estrutura nos dados para identificar uma organização válida e conveniente.

Os métodos tradicionais de agrupamento se decompõem em 2 classes: hierárquico e por partição. O agrupamento por hierarquia produz uma série de subdivisões que se aninham à subdivisão em sequência com base em sua similaridade, produzindo um dendrograma (Jain & Dubes, 1988), (Jain et al., 1999). O agrupamento hierárquico ainda se decompõe em aglomerativo, em que todos os dados são *clusters* que vão se aglomerando a cada iteração para formar os agrupamentos finais, e o divisivo, que se inicia em um único grande grupo e vai se dissolvendo em conjuntos menores, (Berkhin, 2006), (Gan et al., 2007). Ao contrário da técnica aludida, o agrupamento por partição produz uma única repartição, num esforço de manter os grupos naturais dispostos nos dados (Jain & Dubes, 1988). A Figura 14 apresenta o diagrama dos métodos convencionais de aglomeração.

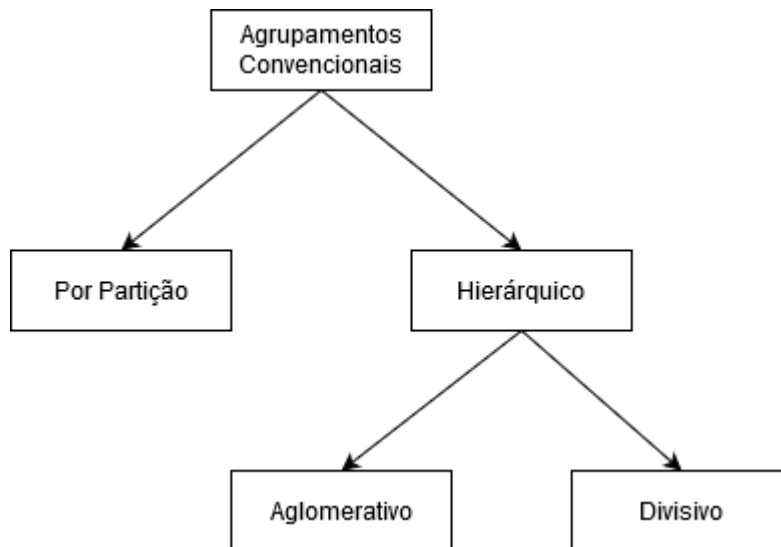


Figura 14 - Diagrama de métodos de *clustering* convencionais. Adaptado de Gan (2007).

A técnica de *clustering* pode ser feita utilizando uma RNA. Para tanto, é necessário um aprendizado não-supervisionado. Isso significa que a rede não recebe nenhuma orientação, ou seja, somente o conjunto de entradas é fornecido, não existindo uma saída pré-definida (Berkin, 2006), (Jain et al., 1999), (Silva et al., 2016). Neste trabalho são usadas ambas técnicas de agrupamento para obter melhor resultado no desenvolvimento prático do trabalho. Dessa forma, foram escolhidos, com base na literatura, os Mapas Auto-Organizáveis de Kohonen para a técnica por partição e para a hierarquia é utilizado o processo aglomerativo.

4.2. Mapas Auto-Organizáveis de Kohonen

O Mapa Auto Organizável de Kohonen é atrativo para a análise de *clusters* por apresentar características de visualização rápida (Affonso, 2011), (Gan et al., 2007), (Vesanto & Alhoniemi, 2000), fornecendo uma maneira fácil de verificar a relação das variáveis para o agrupamento dos dados. O SOM, desenvolvido por Kohonen (1995), é útil comparado às outras redes por ser capaz de representar um conjunto de dados multidimensional em um espaço bidimensional (Affonso, 2011), (Gan et al., 2007), (Kohonen, 1995), (Silva et al., 2016).

A arquitetura dessa rede é composta por 2 camadas: uma camada de entrada e uma camada de Kohonen, organizada em 2 dimensões. A quantidade de neurônios da camada de entrada equivale ao número de propriedades de cada amostra. Bem como cada neurônio dessa camada tem uma conexão *feedforward* com cada neurônio da camada de Kohonen (Gan et al., 2007). Os Mapas Auto-Organizáveis de Kohonen se ordenam por meio de treinamentos

competitivos, tendo como princípio a competição entre os neurônios, em que o vencedor tem seus pesos modificados para se adequar ao próximo vetor de entrada (Silva et al., 2016). A definição do neurônio vencedor se dá pela proximidade entre o vetor de entrada e o vetor de peso do neurônio (Affonso, 2011), (Silva et al., 2016). Essa distância é concebida pela distância Euclidiana, expressa pela Equação (4.1).

$$i(x) = \arg \min \|x - w_j\| \quad j = 1, 2, \dots, n \quad (4.1)$$

Onde x é o vetor de entrada e w_j é o vetor de peso.

Ao obter um vencedor, já tendo arranjado os neurônios no mapa topológico, deve-se definir um critério de vizinhança entre os neurônios para que esse campeão se encontre no centro dos nós ao redor, a fim de que haja um ajuste tanto do vencedor quanto da vizinhança (Affonso, 2011), (Kohonen, 1995). Para que os vizinhos mais afastados do neurônio vitorioso tenham um valor de ajuste menor do que o campeão, é aplicada a função do tipo Gaussiana na vizinhança (Silva et al., 2016), produzindo uma suavização nos vetores de peso desses neurônios (Kohonen, 1995). Assim, a Equação (4.2) se aplica ao vencedor e a Equação (4.3) é utilizada na vizinhança.

$$w_{atual}^{(v)} = w_{anterior}^{(v)} + \eta \cdot (x^k - w_{anterior}^{(v)}) \quad (4.2)$$

$$w_{atual}^{(vizinho)} = w_{anterior}^{(vizinho)} + \eta \cdot \alpha^{(vizinho)} \cdot (x^k - w_{anterior}^{(vizinho)}) \quad (4.3)$$

Onde $\alpha^{(vizinho)}$ é o operador Gaussiano, dado pela expressão:

$$\alpha^{(vizinho)} = e^{-\frac{\|w_{anterior}^{(v)} - w_{anterior}^{(vizinho)}\|^2}{2\sigma^2}} \quad (4.4)$$

De acordo com Silva et al. (2016), o aprendizado competitivo se assemelha ao *k-means* por ter o mesmo critério de proximidade entre as amostras e neurônios, porém no método *k-means* é necessário especificar a quantidade de grupos para classificar os objetos a priori.

4.3. Clustering Hierárquico

De acordo com o que foi dito no começo desse capítulo, o agrupamento hierárquico divide os dados em uma sequência de partições que vão se aninhando de acordo com o critério de similaridade. Esse agrupamento se divide em aglomerativo, em que cada objeto é um *cluster* e a cada iteração ocorre a união com outros objetos semelhantes até formar um único agrupamento; e divisivo, que se inicia em um único grande agrupamento contendo todas as amostras e recursivamente se divide em conjuntos menores (Berkhin, 2006). A Figura 15 demonstra como funciona o processo de *clustering* utilizando o método hierárquico.

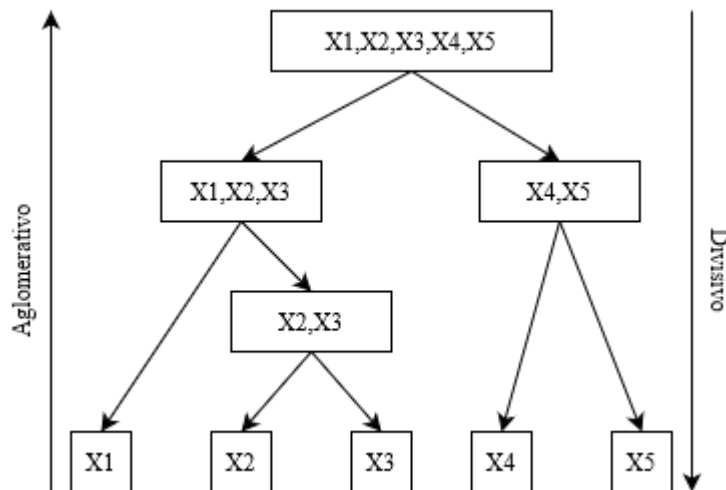


Figura 15 – Métodos de agrupamento hierárquico. Adaptado de Gan (2007).

Esse algoritmo é representado por um dendrograma, em que cada nó configura um grupo de dados semelhantes (Berkhin, 2006), (Jain et al., 1999), que satisfaz a condição

$$h(A) \leq h(B) \Leftrightarrow A \subseteq B \quad (4.5)$$

para todos os subconjuntos dos pontos de dados A e B, em que $h(A)$ e $h(B)$ denotam as alturas de A e B, respectivamente (Gan et al., 2007). A Figura 16 apresenta o dendrograma que funciona através dessa condição.

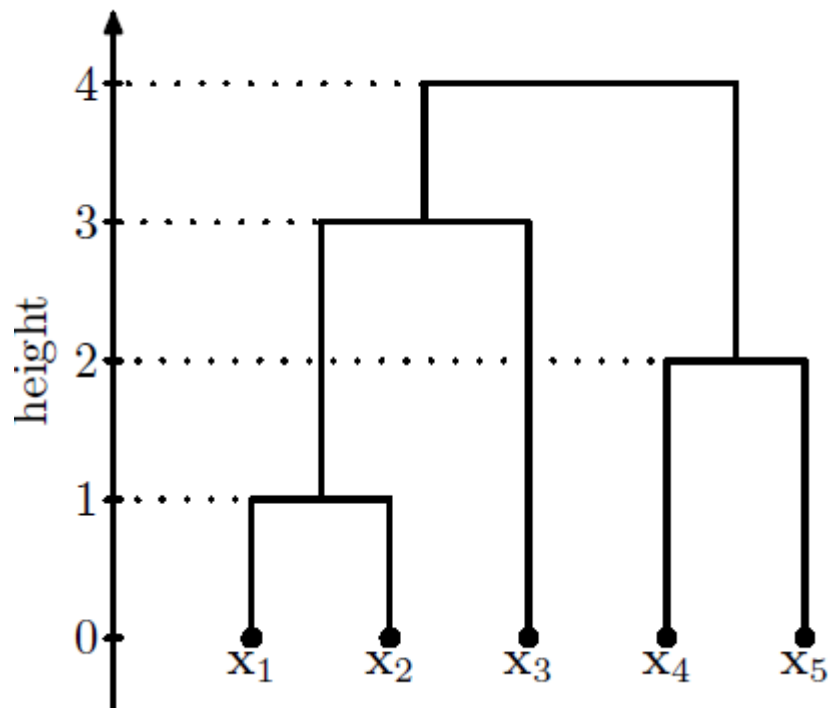


Figura 16 – Dendrograma de 5 elementos. Fonte: Gan (2007).

Para achar a (dis)similaridade entre cada par de objetos da base de dados, calcula-se a distância entre essas observações (Gan et al., 2007), dada pela distância Euclidiana, já expressa na Equação (4.1). Após esse cálculo, é possível determinar como os objetos no conjunto de dados devem ser agrupados em *clusters* através da métrica de ligação (Gan et al., 2007), (Jain et al., 1999). No método aglomerativo, ao ter informações a respeito da distância entre os dados, a função de ligação associa pares desses dados a *clusters* compostos por 2 elementos que estão perto desses objetos (Mathworks, s.d.). Assim, esses novos *clusters* se agrupam com outros objetos criando grupos ainda maiores através dessa mesma função.

A métrica de ligação pode ser completa, única, média, centroide, mediana, Ward ou ponderada. A primeira função usa a maior distância entre objetos nos 2 *clusters*. De forma inversa, a ligação única utiliza a menor distância entre elementos em pares de *clusters* diferentes. A média usa a distância média entre todos os pares de observações em 2 *clusters* distintos, a ligação por centroide usa a distância Euclidiana entre os centros de 2 aglomerados, a função mediana usa a distância Euclidiana entre os centroides ponderados de 2 grupos. Já a ligação de Ward faz a soma incremental dentro do *cluster* como resultado da junção de 2 *clusters* e, por fim, a média ponderada usa uma definição recursiva para a distância entre 2 agrupamentos (Gan et al., 2007), (Jain et al., 1999), (Mathworks, s.d.).

Capítulo 5: Desenvolvimento Prático dos Métodos de Agrupamento

Esse tópico é destinado a descrever a metodologia implantada no agrupamento de dados. São apresentados, também, os resultados e conclusão.

5.1. Metodologia

Assim como na investigação utilizando a análise estatística descritiva, as amostras são compostas por 830 indivíduos pronunciando as vogais /a/, /i/ e /u/ sustentadas nos tons alto, baixo e normal, resultando numa base de dados com 7470 amostras. Os dados contêm 7 variáveis e são organizados em grupo de Controlo, Carcinoma, Cisto, Disfonia Espasmódica, Disfonia Funcional, Disfonia Hiperfuncional, Disfonia Hipofuncional, Disfonia Hipotônica, Disfonia Psicogênica, Edema de Reinke, Granuloma, Granuloma por Intubação, Laringite Crônica, Paralisia das Cordas Vocais, Pólipos das Cordas Vocais, Tumor na Hipofaringe e Tumor na Laringe. Esses grupos ainda se ordenam em gênero feminino seguido do sexo masculino. Esses 7 parâmetros foram colhidos da base de dados implementada por Fernandes (2018), que os extraiu do SVD. Esses atributos são, nomeadamente, o *jitter* absoluto, *jitter* relativo, *shimmer* absoluto, *shimmer* relativo, HNR, NHR e autocorrelação.

Os dados brutos precisam ser preparados para que a análise de *clusters* seja feita (Gan et al., 2007). Portanto, aplica-se uma normalização que varia entre 0 a 1 para os atributos, onde identifica-se os valores máximo e mínimo de cada parâmetro. Para cada fator obtém-se a diferença dessas medidas e a diferença entre todas as amostras desse atributo pelo menor valor extraído dessa propriedade. A partir dessas subtrações, aplica-se a razão entre a segunda e primeira subtração supracitadas, alcançando o valor normalizado da amostra. A Equação (5.1) mostra essa expressão, em que i é o parâmetro a ser normalizado.

$$X_i = \frac{X_i - X_{i \max}}{X_{i \max} - X_{i \min}} \quad i = 1, 2, \dots, 7 \quad (5.1)$$

Foram desenvolvidos 2 algoritmos com diferentes abordagens para o agrupamento. O primeiro algoritmo trabalha com *clustering* por partição, mais especificamente os Mapas Auto-Organizáveis de Kohonen, enquanto o segundo algoritmo utiliza *clustering* hierárquico aglomerativo.

5.2. Mapas Auto-Organizáveis de Kohonen

O algoritmo do SOM foi executado no *software* MATLAB® através de uma *toolbox* desenvolvida pelo próprio programa. As condições da rede foram ajustadas empiricamente de tal forma que ela apresenta uma dimensão de 144 neurônios. O tipo de grade utilizado na rede foi a forma hexagonal, por ser mais eficaz para visualização (Kohonen, 1995). Como Kohonen (1995) sugere, o raio inicial do bairro pode ser até metade do diâmetro da rede, isto porque, caso o tamanho desse raio se inicie muito pequeno ou muito grande a ordenação dos neurônios não ocorre de forma efetiva. Desse modo, o raio inicial de tamanho 10 foi escolhido experimentalmente.

A fim de obter uma boa precisão estatística, a fase de ordenação deve ser relativamente curta, à medida que fase final de convergência deve ser grande, para que haja uma boa precisão estatística. Logo, a rede apresenta 72000 épocas, onde a fase de ordenação contém apenas 2000 épocas.

5.3. Clustering Hierárquico Aglomerativo

Assim como o SOM, o aglomeramento hierárquico também foi desenvolvido no *software* MATLAB® por meio do *Statistics and Machine Learning Toolbox*. A princípio, utiliza-se a equação Euclidiana, a fim de determinar a distância entre cada objeto da base de dados. Em sequência, emprega-se a métrica de ligação para criar os *clusters*. Como Gan (2007) afirma, essa função pode ser desvantajosa, pois, ao variar a medida de similaridade entre *clusters*, os resultados podem variar. No entanto, utilizando essa medida é possível criar a árvore hierárquica. No dendrograma, a ligação entre as amostras é representada pelas linhas que formam a letra U de cabeça para baixo (Mathworks, s.d.) e a altura dada para essas linhas retrata a distância entre os objetos.

Para verificar se a medida de ligação foi escolhida corretamente, basta comparar a distância copenética, ou seja, a altura entre 2 objetos na árvore, com a distância obtida pela equação Euclidiana. Para um resultado consistente, essa correlação deve apresentar um valor próximo de 1, comprovando que a solução desse *cluster* representa os dados originais (Mathworks, s.d.). Para tanto, ao analisar as funções de vinculação com os dados originais, conclui-se que a ligação por centroide é a mais apropriada, por apresentar um valor de coeficiente superior ao restante, apresentando uma correlação copenética de 0,8966. A Tabela 11 mostra a relação entre as métricas de ligação e suas distâncias copenéticas.

Tabela 11 – Relação entre métricas de ligação e suas distâncias copenéticas.

Métricas de Ligação	Distância Copenética
Média	0,8771
Centroide	0,8966
Completo	0,8419
Mediana	0,8602
Ward	0,7431
Média Ponderada	0,8072
Única	0,8508

Com a árvore criada, é possível particionar os dados para criar os agrupamentos. Para tanto, tendo como princípio os resultados da análise estatística descritiva, que dividiu os dados em 2 conjuntos, utilizou-se esse valor para especificar, nesse algoritmo, a quantidade de grupos a serem criados. Dessa forma, é possível saber quais amostras pertencem a quais *clusters*.

5.4. Resultados e Discussões

Ao implementar o algoritmo SOM na base de dados é possível averiguar a quantidade de agrupamentos por meio das distâncias entre os neurônios vizinhos. A Figura 17 apresenta essa distância entre a vizinhança, em que as cores mais escuras representam distâncias maiores, enquanto as cores mais claras representam distâncias menores.

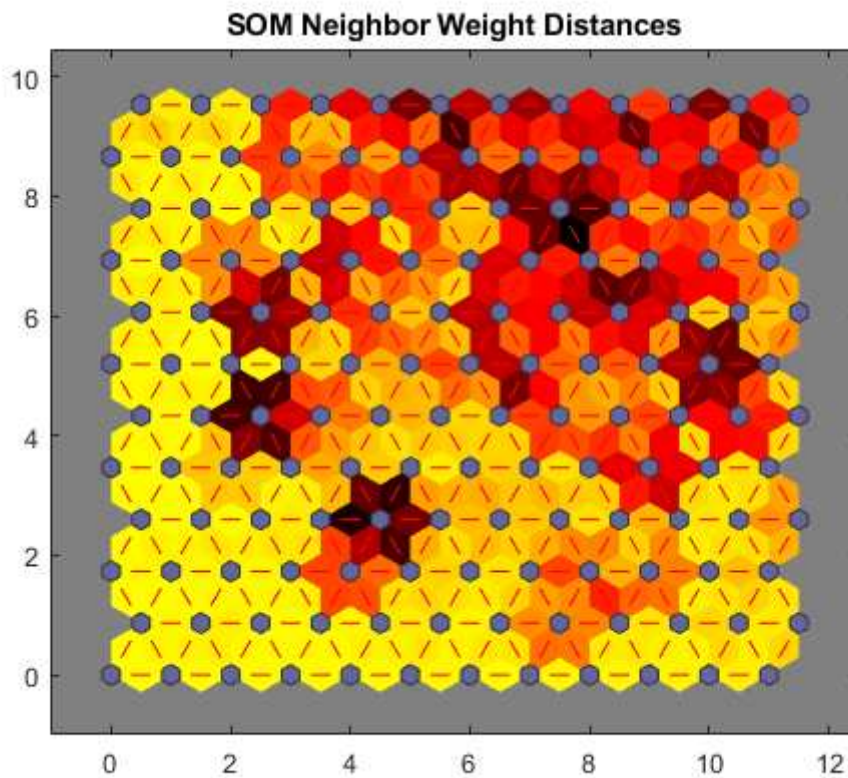


Figura 17 - Distância entre a vizinhança dos neurônios.

Um grande grupo de cores claras aparece na região inferior esquerda, implicando num grupo de pontos de dados fortemente agrupados. No canto superior direito estão os segmentos mais escuros, indicando que esses dados estão mais afastados um dos outros. Em vista disso, é possível concluir que existem 2 *clusters*, sendo que em um deles seus elementos estão bem próximos, enquanto no segundo conjunto os objetos estão mais dispersos.

A fim confirmar a veracidade dessa separação em 2 conjuntos, o algoritmo de *clustering* hierárquico aglomerativo é implementado. Com a função de vinculação centroide é possível obter a Figura 18, que apresenta o dendrograma desse método. Na árvore hierárquica nota-se 2 conjuntos aglomerados, um com uma grande quantidade de elementos, e outro grupo menor. O retângulo amarelo, disposto nessa figura, indica esse aglomeramento maior, ao passo que a cor vermelha representa o menor conjunto agrupado.

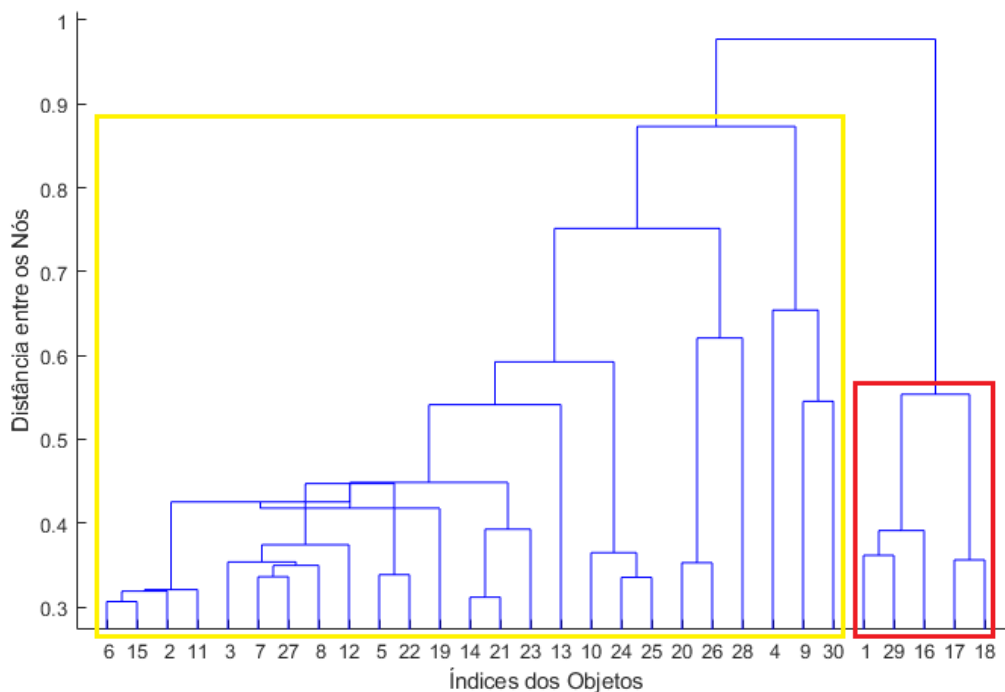


Figura 18 – Árvore hierárquica aglomerativa dividida em 2 conjuntos.

Quando são apurados os valores de cada grupo disposto nos *clusters* hierárquicos, constatou-se que os grupos se dividiram em 7183 amostras pertencentes a um conjunto e 287 a outro. No maior grupo estão inteiramente dispostos os conjuntos Cisto, Disfonia Hipotônica, Disfonia Hipofuncional e Granuloma. No conjunto com 287 amostras estão inseridos 112 sujeitos que se dividem em 2 indivíduos do grupo de Controle, 7 do Carcinoma, 20 da Disfonia Espasmódica, 6 da Disfonia Funcional, 11 da Disfonia Hiperfuncional, 7 da Disfonia Psicogênica, 8 do Edema de Reinke, 1 do Granuloma por Intubação, 5 da Laringite Crônica, 37 da Paralisia nas Cordas Vocais, 5 do Pólipo nas Cordas Vocais, 1 do Tumor na Hipofaringe e 2 do Tumor na Laringe. Essa divergência entre valores acontece, pois, cada indivíduo possui 9 amostras (vogais e tons) e muitos dos sujeitos pertencentes ao menor agrupamento tem apenas 1 desses 9 dados inseridos nesse menor grupo.

Com intuito de confirmar a veracidade da quantidade de *clusters* existentes nos dados, utilizou-se a função métrica Davies-Bouldin disposta no MATLAB®. Aplicou-se o algoritmo de *k-means* ao mapa com valores de agrupamentos divergentes do resultado esperado. No entanto, ao avaliar as 7470 observações, esse índice retornou um valor otimizado de 2 *clusters*.

Para averiguar a estrutura da base de dados, faz-se um novo agrupamento, onde utiliza-se como base o resultado do Diagrama de Caixa no Capítulo 3. Assim, as patologias que não se agrupam a nenhum conjunto na análise estatística descritiva são retiradas, mais especificamente

o Carcinoma, o Granuloma e o Tumor na Laringe. Logo, essa nova base de dados contém 7245 observações, ou seja, 805 indivíduos. Aplica-se os métodos de *clustering* com as mesmas configurações para os grupos restantes, sendo eles o grupo de Controlo, Cisto, Disfonia Espasmódica, Disfonia Funcional, Disfonia Hiperfuncional, Disfonia Hipofuncional, Disfonia Hipotônica, Disfonia Psicogênica, Edema de Reinke, Granuloma por Intubação, Laringite Crônica, Paralisia das Cordas Vocais, Pólipos das Cordas Vocais e Tumor na Hipofaringe. A Figura 19 e a Figura 20 apresentam a distância entre a vizinhança do SOM e o dendrograma do *clustering* hierárquico aglomerativo, respectivamente.

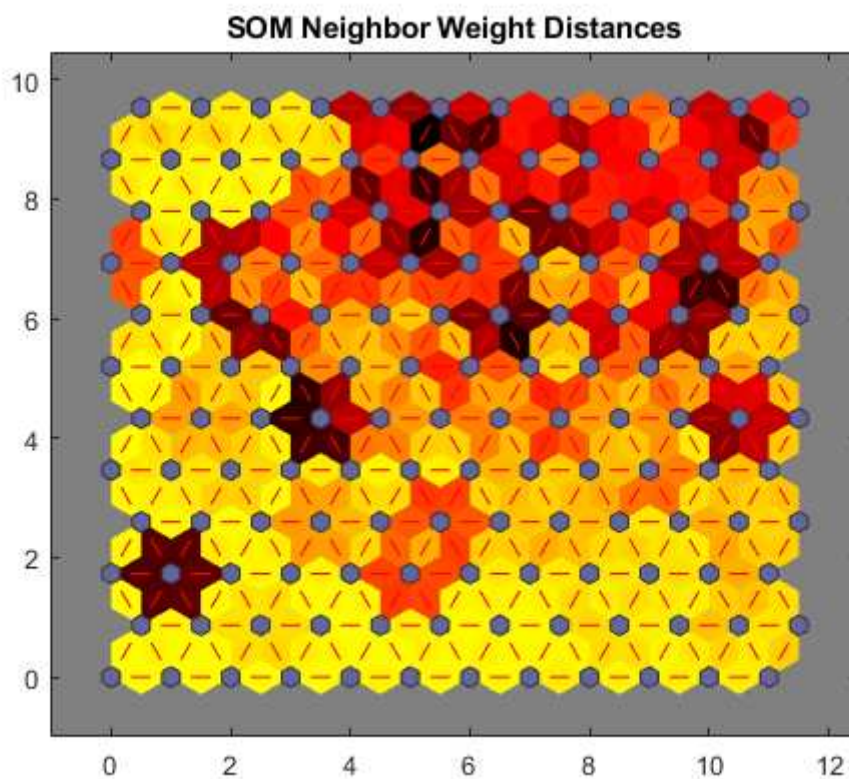


Figura 19 – Distância entre a vizinhança dos neurônios com o novo conjunto.

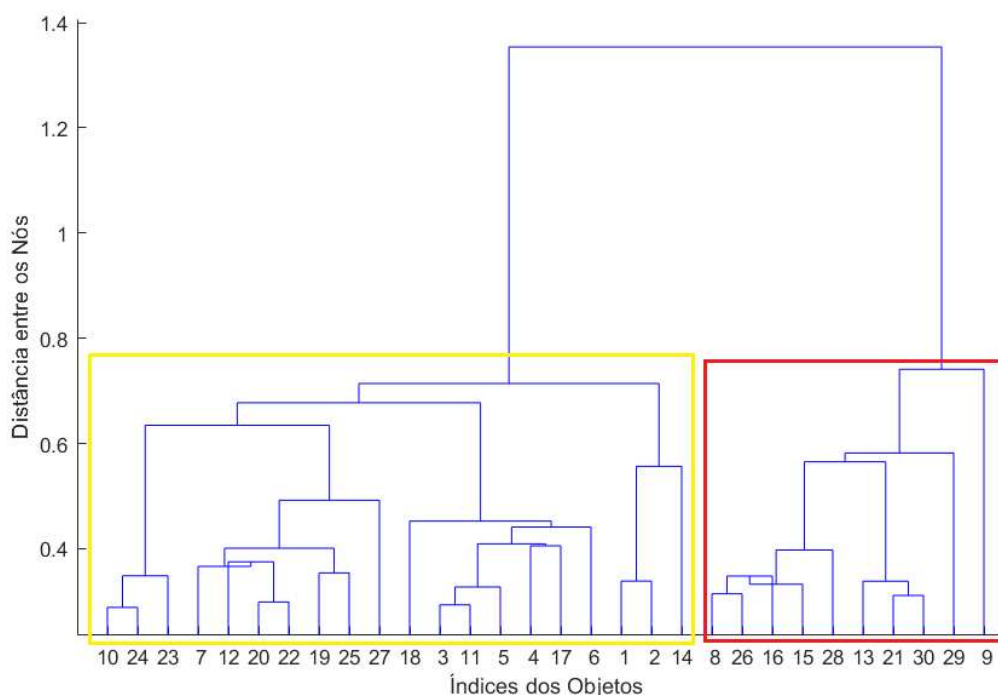


Figura 20 - Árvore hierárquica aglomerativa com novo agrupamento.

Assim como na Figura 17, a distância entre as vizinhanças do neurônios que o SOM apresenta na Figura 19 contém 2 conjuntos, sendo um deles com cores claras na região inferior à esquerda e outro com tons mais escuros no canto superior à direita. Nessa divisão entre esses agrupamentos não há uma separação precisa, onde pode ser visto que no centro da figura os tons se mesclam.

A mesma coisa acontece em relação a Figura 18 e Figura 20, com o dendrograma da base de dados completa e a sem 3 patologias, respectivamente. Na Figura 20, os conjuntos também se separam em 2 agrupamentos, onde o retângulo amarelo implica no maior aglomeramento, enquanto a cor vermelha representa o menor conjunto agrupado. O primeiro agrupamento apresenta a maior quantidade de amostras, somando 7207 elementos. Nesse grupo estão inseridos completamente o grupo de Controle, o Cisto, a Disfonia Hipofuncional, a Disfonia Hipotônica, a Disfonia Psicogênica, o Granuloma por Intubação, o Pólipo das Cordas Vocais e o Tumor na Hipofaringe. No menor grupo estão presentes 38 amostras dispostas em 20 indivíduos, sendo eles 2 sujeitos da Disfonia Espasmódica, 1 da Disfonia Funcional, 4 da Disfonia Hiperfuncional, 1 do Edema de Reinke, 2 da Laringite Crônica e 10 da Paralisia nas Pregas Vocais.

Como na maior base de dados, a métrica Davies-Bouldin também foi implementada com o algoritmo de *clustering k-means* na nova base de dados, obtendo também um valor otimizado de 2 *clusters*.

5.5. Conclusão

Ao confrontar os resultados de ambos algoritmos de agrupamento aplicados ao mesmo banco de dados e utilizando a mesma medida de proximidade (distância Euclidiana), constata-se que ambos apresentam o mesmo resultado. São observados 2 conjuntos, sendo que um deles apresenta uma quantidade maior de objetos presentes, enquanto o outro possui uma pequena quantidade de dados. Além disso, pode-se perceber que os grupos Cisto, Disfonia Hipotônica, Disfonia Hipofuncional e Granuloma pertencem completamente ao maior conjunto.

O Mapa Auto Organizável de Kohonen, apesar de possuir técnicas aprimoradas de visualização, pode apresentar certa dificuldade de interpretação de resultados, já que os recursos responsáveis por uma atribuição específica de *cluster* não são evidentes na representação disposta no mapa (Callan, Kent, Roy & Tasko, 1999). Logo, fez-se necessário um novo método de agrupamento para entender melhor o que acontece com os dados. Contudo, essa nova técnica pode apresentar a desvantagem de apresentar resultados diferentes dependendo do método de vinculação utilizado. Para isso, foi utilizada a função de ligação por centroide, com maior correlação copenética de 0,8966.

Ao implementar os mesmos métodos de *clustering* para um novo conjunto sem as patologias Carcinoma, Granuloma e Tumor na Laringe, obteve-se os mesmos resultados que a implementação feita para a maior base de dados. Portanto, como conclusão final, pode-se afirmar que essa separação em 2 conjuntos pode ser considerada aceitável, posto que, apesar das divergências e desvantagens que ambos métodos apresentam, eles obtiveram o mesmo resultado para as 2 bases de dados. No entanto, a Figura 17 e a Figura 19 apresentam uma mescla entre as cores que definem os 2 grupos bem ao centro da imagem. No *clustering* aglomerativo com a base de dados completa, 52 sujeitos do menor conjunto possuem apenas 1 das 9 amostras pertencentes a este indivíduo e na base de dados reduzida, 18 indivíduos do menor conjunto possuem apenas 1 das 9 amostras pertencentes a este grupo. Isto pode implicar numa divisão errônea entre os conjuntos por parte de ambos os métodos. Apesar disso, a

avaliação da quantidade de agrupamentos utilizando o índice de Davies-Bouldin também confirma a quantidade de 2 *clusters* existentes nos dados (Davies & Bouldin, 1979).

Capítulo 6: Conclusão e Trabalhos Futuros

Neste capítulo serão apresentadas as considerações finais por meio de uma discussão concisa. Também serão discutidos a conclusão e os trabalhos futuros.

6.1. Conclusão Geral

A intenção desta dissertação foi aplicar técnicas de agrupamento de sujeitos dos grupos de Controlo, Carcinoma, Cisto, Disfonia Espasmódica, Disfonia Funcional, Disfonia Hiperfuncional, Disfonia Hipofuncional, Disfonia Hipotônica, Disfonia Psicogênica, Edema de Reinke, Granuloma, Granuloma por Intubação, Laringite Crônica, Paralisia das Cordas Vocais, Pólipos das Cordas Vocais, Tumor na Hipofaringe e Tumor na Laringe por meio das perturbações em frequência e amplitude, razões entre harmônico-ruído e autocorrelação. Isto para entender como se estrutura os sinais de voz, já que estes são complexos.

Este estudo apresenta suas colaborações ramificadas em 2 categorias: a averiguação de similaridades por meio da análise estatística descritiva, disposta no Capítulo 3, e o agrupamento de dados, descrito no Capítulo 5.

A análise estatística descritiva feita através do Diagrama de Caixa verificou a semelhança entre os dados por meio da comparação entre cada grupo pré-estabelecido pelo SVD. Dessa forma, os conjuntos foram contrapostos entre si para examinar qual das 3 situações mais se adequa aos pares analisados. Essas 3 circunstâncias são: o agrupamento, o possível não agrupamento e o não agrupamento entre as caixas.

Assim como no artigo apresentado por Oliveira et al. (2020), essa investigação observou-se que 3 patologias, o Carcinoma, o Granuloma e o Tumor na Laringe, se diferem do restante das amostras, indicando que estas não se agrupam a nada. De forma inversa, o Tumor na Hipofaringe, a Disfonia Hipotônica, a Disfonia Hipofuncional e Disfonia Funcional podem se unir a maioria dos outros grupos.

No agrupamento dos dados, foram utilizadas 2 redes neuronais artificiais não-supervisionadas com abordagens divergentes. A primeira RNA trabalhou com os Mapas Auto-Organizáveis de Kohonen, ao passo que a segunda RNA utilizou o agrupamento hierárquico

aglomerativo. A implementação desses algoritmos empregou a mesma medida de proximidade, sendo a distância Euclidiana.

Ambos os métodos de *clustering* apresentam desvantagens, seja pela dificuldade de interpretação de resultados ou por apresentar divergência nesses resultados quando um dado é alterado. Contudo, essas 2 técnicas apresentaram o mesmo resultado, onde se formaram 2 grupos, sendo visível que um deles apresenta uma quantidade maior de objetos comparado ao outro. Além disso, um método de avaliação de *clusters* foi empregado, resultando no mesmo número de agrupamentos encontrados. A fim de confirmar a estrutura dos dados, fez-se um novo estudo desses métodos de *clustering*, retirando os conjuntos que, de acordo com o Diagrama de Caixa, não se agrupam a nenhum outro, sendo eles o Carcinoma, o Granuloma e o Tumor na Laringe. O resultado dessa implementação foi semelhante aos resultados apresentados pela base de dados completa, ou seja, formaram-se 2 conjuntos.

De acordo com as informações identificadas por ambas as categorias utilizadas para entender a estrutura dos dados constata-se que, apesar das técnicas verificarem a mesma quantidade de conjuntos, o Diagrama de Caixa e os métodos de *clustering* agrupam de maneira diferente a base de dados. Portanto, esse estudo prova que a base de dados implementada por Fernandes (2018) utilizada nas 2 categorias, sendo elas a análise estatística descritiva e os métodos de *clustering*, não se complementam, por fornecerem informações divergentes a respeito da formação dos *clusters*. Por conseguinte, nota-se que, apesar de serem uma ferramenta poderosa, os métodos de *clustering* não foram capazes de encontrar estrutura nos dados para identificar uma organização válida e conveniente.

Vários obstáculos surgiram ao longo deste trabalho, tanto no conhecimento a respeito do tema quanto no desenvolvimento da dissertação em si. Contudo, o objetivo proposto no início desse trajeto foi atingido, mesmo que diferente da forma pretendida. A escolha dos parâmetros foi relativamente fácil, por fazer parte de um ramo já estudado. O Diagrama de Caixa foi escolhido também por já ter sido trabalhado ao longo do caminho acadêmico. A decisão a respeito dos métodos de *clustering* foi mais complexa, por haver algumas opções, mas não ter sido estudada anteriormente. Essa escolha foi feita com base nas leituras a respeito de métodos de agrupamento e, no momento, ambos os métodos selecionados pareceram mais convenientes. O percurso do início ao fim desta dissertação foi dificultoso, porém, extremamente gratificante por proporcionar um alicerce para um futuro doutorado na área de *clustering*, que acabou sendo um campo de grande interesse.

6.2. Trabalhos Futuros

Para trabalhos futuros, é sugerida a implementação dessas mesmas redes neuronais utilizando outros parâmetros extraídos do sinal da voz, tal qual o *mel-frequency cepstral coefficient* (MFCC), *perceptual linear prediction* (PLP), *linear prediction cepstral coefficient* (LPCC), entre outros. Ou mesmo aplicar transformadas, como a de Fourier ou Wavellet, nos sinais vocais, a fim de obter outros atributos.

Além do mais, pode-se aplicar outras ferramentas de análise de *clustering*, como *k-means*, *Gaussian Mixture Models* (GMM), *K-Nearest Neighbors* (KNN) na base de dados a fim de obter novos agrupamentos que complementem o Diagrama de Caixa.

Referências

- Affonso, G. S., 2011. Mapas Auto-Organizáveis De Kohonen (SOM) Aplicados Na Avaliação De Parâmetros Da Qualidade Da Água. Mestre. Autarquia associada à Universidade de São Paulo.
- Altman, K. W., 2007. Vocal Fold Masses. *Otolaryngologic Clinics of North America*, 40(5), pp.1091-1108.
- Berkhin, Pavel. (2002). Survey Of Clustering Data Mining Techniques. A Survey of Clustering Data Mining Techniques. *Grouping Multidimensional Data: Recent Advances in Clustering*.
- Boone, D., McFarlane, S., Von Berg, S. & Zraick, R., 2014. *The Voice And Voice Therapy*. 9th ed. EUA: Pearson, pp.9-14.
- British Medical Journal. Post-Intubation Granuloma, v. 4, n. 5888, p. 313-314, nov./1973.
- Callan, D., Kent, R., Roy, N. & Tasko, S., 1999. Self-Organizing Map for the Classification of Normal and Disordered Female Voices. *Journal of Speech, Language, and Hearing Research*, 42(2), pp.355-366.
- Chen, H., Jen, Y., Wang, C., Lee, J. & Lin, Y., 2007. Etiology of Vocal Cord Paralysis. *ORL*, (3), pp.167-171.
- Cleveland Clinic. Hypopharyngeal Cancer. Disponível em: <https://my.clevelandclinic.org/health/diseases/12181-hypopharyngeal-cancer>. Acesso em: 05 dez. 2019.
- Cleveland Clinic. Vocal Cord Lesions (Nodules, Polyps and Cysts). Disponível em: <https://my.clevelandclinic.org/health/diseases/15424-vocal-cord-lesions-nodules-polyps-and-cysts>. Acesso em: 29 nov. 2019.
- Clínica San Vicente. Voice Disorders: DYSPHONIA. Disponível em: <https://www.clinicasanvicente.es/en/servicios/voice-disorders-dysphonia/>. Acesso em: 29 nov. 2019.
- Columbia University. Reinke's Edema. Disponível em: <https://www.entcolumbia.org/health-library-0/reinkes-edema>. Acesso em: 30 nov. 2019.

- Cordeiro, H., 2016. Reconhecimento De Patologias Da Voz Usando Técnicas De Processamento Da Fala. Doutor. Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa.
- Davies, D. & Bouldin, D., 1979. A Cluster Separation Measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-1(2), pp.224-227.
- Dr. Wohlt. Functional Dysphonia. Disponível em: <https://www.wohlt.com/functional-dysphonia/>. Acesso em: 05 dez. 2019.
- Fawcus, M., 1991. *Voice Disorders And Their Management*. 2 ed. Londres: Springer-Science+Business Media, B.V., pp.1-392.
- Felippe, A., Grillo, M. & Grechi, T., 2006. Normatização de medidas acústicas para vozes normais. *Revista Brasileira de Otorrinolaringologia*, 72(5), pp.659-664.
- Fernandes, J., 2018. Determinação Da Autocorrelação, HNR E NHR Para Análise Acústica Vocal. Mestre. Instituto Politécnico de Bragança.
- Finger, L., Cielo, C. & Schwarz, K., 2009. Medidas vocais acústicas de mulheres sem queixas de voz e com laringe normal. *Revista Brasileira de Otorrinolaringologia*, 75(3), pp.432-440.
- Gan, G., Ma, C. & Wu, J., 2007. *Data Clustering*. 1 ed. Philadelphia, Pa.: SIAM, Society for Industrial and Applied Mathematics, pp.1-487.
- Hillman, R., Holmberg, E., Perkell, J., Walsh, M. & Vaughan, C., 1989. Objective Assessment of Vocal Hyperfunction. *Journal of Speech, Language, and Hearing Research*, 32(2), pp.373-392.
- Holinger, L., Holinger, P. & Holinger, P., 1976. Etiology of Bilateral Abductor Vocal Cord Paralysis. *Annals of Otology, Rhinology & Laryngology*, 85(4), pp.428-436.
- Holmberg, E., Doyle, P., Perkell, J., Hammarberg, B. & Hillman, R., 2003. Aerodynamic and acoustic voice measurements of patients with vocal nodules: variation in baseline and changes across voice therapy. *Journal of Voice*, 17(3), pp.269-282.
- Jain, A., Murty, M. & Flynn, P., 1999. Data Clustering: A Review. *ACM Computing Surveys*, 31(3), pp.264-323.

- Jain, A. & Dubes, R., 1988. Algorithms For Clustering Data. 1st ed. Michigan: Prentice Hall, pp.1-320.
- John Hopkins Medicine. Vocal Cord Cancer. Disponível em: <https://www.hopkinsmedicine.org/health/conditions-and-diseases/vocal-cord-cancer>. Acesso em: 29 nov. 2019.
- Kandoğan, T., Koç, M. & Aksoy, G., 2009. Effectiveness of voice therapy in hyperfunctional dysphonia in adult patients. *The Turkish Journal of Ear Nose and Throat*, 19(4), pp.198-202.
- Kohonen, T., 1995. Self-Organizing Maps. 1st ed. Berlin, Heidelberg: Springer-Verlag, pp.1-500.
- LaPine, P., 2008. The Relationship between the Physical Aspects of Voice Production and Optimal Vocal Health. *Music Educators Journal*, 94(3), pp.24-29.
- Leinonen, L., Kangas, J., Torkkola, K. & Juvas, A., 1992. Dysphonia Detected by Pattern Recognition of Spectral Composition. *Journal of Speech, Language, and Hearing Research*, 35(2), pp.287-295.
- Leinonen, L., Hiltunen, T., Laakso, M., Rihkanen, H. & Poppius, H., 1997. Categorization of Voice Disorders with Six Perceptual Dimensions. *Folia Phoniatria et Logopaedica*, 49(1), pp.9-20.
- Mann, P., 2010. Introductory Statistics. 7 ed. EUA: John Wiley & Sons, pp.1-623.
- Martínez, D., Lleida, E., Ortega, A., Miguel, A. & Villalba, J., 2012. Voice Pathology Detection on the Saarbrücken Voice Database with Calibration and Fusion of Scores Using MultiFocal Toolkit. *Communications in Computer and Information Science*, 328(1), pp.99-109.
- Martins, R., Tavares, E., Ranalli, P., Branco, A. & Pessin, A., 2014. Psychogenic dysphonia: diversity of clinical and vocal manifestations in a case series. *Brazilian Journal of Otorhinolaryngology*, 80(6), pp.497-502.
- Mathworks. Hierarchical Clustering. Disponível em: <https://www.mathworks.com/help/stats/hierarchical-clustering.html>. Acesso em: 21 abr. 2020.
- Mayo Clinic. Laryngitis. Disponível em: <https://www.mayoclinic.org/diseases-conditions/laryngitis/symptoms-causes/syc-20374262?p=1>. Acesso em: 04 dez. 2019.

- Medscape. Functional Voice Disorders. Disponível em: <https://emedicine.medscape.com/article/865191-overview> . Acesso em: 11 abr. 2020.
- Michigan Medicine. Functional Dysphonia. Disponível em: <https://www.uofmhealth.org/conditions-treatments/ear-nose-throat/functional-dysphonia>. Acesso em: 02 dez. 2019.
- Morettin, P. & Bussab, W., 2020. Estatística Básica. 7 ed. São Paulo: Saraiva, pp.1-540.
- Morrison, M., Nichol, H. & Rammage, L., 1994. The Management Of Voice Disorders. 1st ed. San Diego: Springer-Science+Business Media, B.V., pp.1-265.
- National Institute on Deafness and Other Communication Disorders (NIDCD). Spasmodic Dysphonia. Disponível em: <https://www.nidcd.nih.gov/health/spasmodic-dysphonia>. Acesso em: 05 dez. 2019.
- National Institute on Deafness and Other Communication Disorders (NIDCD). Taking Care of Your Voice. Disponível em: <https://www.nidcd.nih.gov/health/taking-care-your-voice>. Acesso em: 14 abr. 2020.
- Oliveira, A. A., Dajer, M. E., Fernandes, P. & Teixeira, J.P. (2020). Clustering of Voice Pathologies based on Sustained Voice Parameters. 280-287.
- Panek, D., Skalski, A., Gajda, J. & Tadeusiewicz, R., 2015. Acoustic analysis assessment in speech pathology detection. *International Journal of Applied Mathematics and Computer Science*, 25(3), pp.631-643.
- Parnell, F. & Brandenburg, J., 1970. Vocal Cord Paralysis: A Review of 100 Cases. *The Laryngoscope*, 80(7), pp.1036-1045.
- Pontes, P., Yamasaki, R. & Behlau, M., 2006. Morphological and Functional Aspects of the Senile Larynx. *Folia Phoniatrica et Logopaedica*, 56(3), pp.151-158.
- Pützer, Manfred & Barry, Willian J. Saarbruecken Voice Database: Handbook. Disponível em: <http://stimmdb.coli.uni-saarland.de>. Acesso em: 14 abr. 2020.
- Rihkanen, H., Leinonen, L., Hiltunen, T. & Kangas, J., 1994. Spectral pattern recognition of improved voice quality. *Journal of Voice*, 8(4), pp.320-326.

- Rimoli, C., Martins, R., Catâneo, D., Imamura, R. & Catâneo, A., 2018. Treatment of post-intubation laryngeal granulomas: systematic review and proportional meta-analysis. *Brazilian Journal of Otorhinolaryngology*, 84(6), pp.781-789.
- Rosa, F., Guimarães, I., Pereira, A.L. & Caçador, M. (2011). *A Voz Humana – Comunicação, Arte e Ciência*. Lumen Veritatis.
- Roy, N., 2003. Functional dysphonia. *Current Opinion in Otolaryngology & Head and Neck Surgery*, 11(3), pp.144-148.
- Silbiger, H., 1966. Voice Classification by Hierarchical Clustering. *The Journal of the Acoustical Society of America*, 40(5), pp.1282-1282.
- Silva, I., Spatti, D. & Flauzino, R., 2016. *Redes Neurais Artificiais Para Engenharia E Ciências Aplicadas: Fundamentos Teóricos E Aspectos Práticos*. 2 ed. São Paulo: Artliber, pp.221-241.
- Soltani, M., Ashayeri, H., Modarresi, Y., Salavati, M. & Ghomashchi, H., 2014. Fundamental Frequency Changes of Persian Speakers Across the Life Span. *Journal of Voice*, (3), pp.274-281.
- Teixeira, J. & Fernandes, P., 2014. Jitter, Shimmer and HNR Classification within Gender, Tones and Vowels in Healthy Voices. *Procedia Technology*, 16, pp.1228-1237.
- Teixeira, J., Fernandes, J., Teixeira, F. & Fernandes, P. (2018). Acoustic Analysis of Chronic Laryngitis - Statistical Analysis of Sustained Speech Parameters. 168-175.
- Teixeira, J., Ferreira, D. & Carneiro, S. (2011). Análise acústica vocal - determinação do Jitter e Shimmer para diagnóstico de patologias da fala.
- Teixeira, J., Oliveira, C. & Lopes, C. (2013). Vocal Acoustic Analysis – Jitter, Shimmer and HNR Parameters. *Procedia Technology*. 9.
- Toutouchi, S., Eydi, M., Golzari, S., Ghaffari, M. & Parvizian, N., 2014. Vocal Cord Paralysis and its Etiologies: A Prospective Study. *Journal of Cardiovascular and Thoracic Research*, 6(1), pp.47-50.
- University of Miami Health System. Vocal Fold Nodules, Cysts, and Polyps. Disponível em: [https://umiamihealth.org/en/treatments-and-services/ear,-nose,-and-throat-\(ent\)/laryngology-and-voice/vocal-nodules,-cysts,-and-polyps](https://umiamihealth.org/en/treatments-and-services/ear,-nose,-and-throat-(ent)/laryngology-and-voice/vocal-nodules,-cysts,-and-polyps). Acesso em: 20 nov. 2019.

University of Pittsburgh. Laryngeal Cancer. Disponível em:
<http://www.otolaryngology.pitt.edu/centers-excellence/voice-center/conditions-we-treat/laryngeal-cancer>. Acesso em: 05 dez. 2019.

Vesanto, J. & Alhoniemi, E., 2000. Clustering of the self-organizing map. *IEEE Transactions on Neural Networks*, 11(3), pp.586-600.