



Universidade Federal Da Paraíba

Centro De Tecnologia

Programa de Pós-Graduação em Engenharia Mecânica

Mestrado em Dinâmica e Controle de Sistemas Mecânicos

**DESENVOLVIMENTO DE UM SISTEMA PARA IDENTIFICAÇÃO DE PADRÃO
SONORO DA VOZ HUMANA DE INDIVÍDUOS COM INSUFICIÊNCIA CARDÍACA
UTILIZANDO REDES NEURAIAS ARTIFICIAIS**

por

João Vitor Lira de Carvalho Firmino

Dissertação de Mestrado apresentada à Universidade Federal da Paraíba para obtenção do grau
de Mestre

João Pessoa – Paraíba

Julho, 2021

JOÃO VITOR LIRA DE CARVALHO FIRMINO

**APLICAÇÃO DE REDES NEURAIS ARTIFICIAIS NA IDENTIFICAÇÃO DE
INSUFICIENCIA CARDÍACA UTILIZANDO ANÁLISE SONORA DA VOZ**

Dissertação apresentada ao Curso de Pós-Graduação em Engenharia Mecânica da Universidade Federal da Paraíba, em cumprimento às exigências para obtenção do Grau de Mestre.

Orientador: Professor Dr. Marcelo Cavalcanti Rodrigues
Coorientador: Professor Dr. Marcelo Dantas Tavares de Melo

Catálogo na publicação
Seção de Catalogação e Classificação

F525d Firmino, João Vitor Lira de Carvalho.

Desenvolvimento de um sistema para identificação de padrão sonoro da voz humana de indivíduos com insuficiência cardíaca utilizando redes neurais artificiais / João Vitor Lira de Carvalho Firmino. - João Pessoa, 2021.
132 f. : il.

Orientação: Marcelo Cavalcanti Rodrigues.

Coorientação: Marcelo Dantas Tavares de Melo.

Dissertação (Mestrado) - UFPB/CT.

1. Insuficiência cardíaca. 2. Diagnóstico. 3. Vozes. 4. Análise de sinais. 5. Redes neurais artificiais. I. Rodrigues, Marcelo Cavalcanti. II. Melo, Marcelo Dantas Tavares de. III. Título.

UFPB/BC

CDU 616.12-008.46(043)

**DESENVOLVIMENTO DE UM SISTEMA PARA
IDENTIFICAÇÃO DE PADRÃO SONORO DA VOZ HUMANA
DE INDIVÍDUOS COM INSUFICIÊNCIA CARDÍACA
UTILIZANDO REDES NEURAIS ARTIFICIAIS**

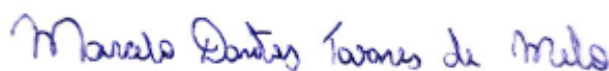
por

JOÃO VITOR LIRA DE CARVALHO FIRMINO

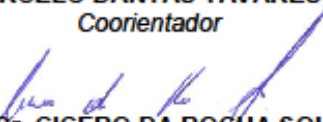
Dissertação aprovada em 22 de julho de 2021



Prof. Dr. MARCELO CAVALCANTI RODRIGUES
Orientador – UFPB



Prof. Dr. MARCELO DANTAS TAVARES DE MELO
Coorientador



Prof. Dr. CICERO DA ROCHA SOUTO
Examinador Interno – UFPB



Prof. Dr. KOJE DANIEL VASCONCELOS MISHINA
Examinador Externo – UFPB

AGRADECIMENTOS

Agradeço a Deus pela saúde e disposição que me permitiu a realização deste trabalho.

Aos meus pais, Antônio Ricardo Firmino e Maria Sonaly Firmino, por todo o amor e apoio recebido para minha formação pessoal e profissional.

Aos meus irmãos Antônio, Arthur, Luiza e Miguel, por todos os momentos compartilhados.

A Lairanne Costa pelo amor e apoio incondicional em todos os momentos.

Aos familiares mais próximos que contribuíram, de alguma maneira, para o meu sucesso acadêmico.

Aos professores Marcelo Cavalcanti Rodrigues e Marcelo Dantas Tavares de Melo, pela confiança, orientação e conhecimentos passados na elaboração desta dissertação.

Ao técnico José Gilsivan Cartaxo da Silva pelos conhecimentos passados para a elaboração desta dissertação.

Aos professores da banca por terem aceitado fazer a avaliação deste trabalho.

A todos os professores que fizeram parte da minha caminhada estudantil, contribuindo com o meu crescimento pessoal e intelectual.

Ao CNPq, pelo apoio financeiro durante a vigência deste curso.

Agradeço também a todos os demais que de alguma forma contribuíram para a realização deste trabalho.

RESUMO

A insuficiência cardíaca é uma doença que incapacita o coração de bombear o sangue, de forma adequada, para nutrir todo o organismo. Atualmente, os principais métodos de diagnóstico dessa patologia são realizados clinicamente por meio da mensuração do peptídeo natriurético do tipo B (BNP). Como as doenças cardiovasculares são as maiores causas de morte prematura, o desenvolvimento de novas tecnologias para a identificação dessas doenças é de grande importância. Assim, o presente trabalho apresenta o desenvolvimento de um aplicativo de identificação das distorções vocais causadas pela insuficiência cardíaca em um indivíduo saudável. Para a elaboração do software, foram coletadas as vozes de 142 indivíduos, separados por sexo, idade. Dessas 142, 84 vozes de pessoas já diagnosticadas com insuficiência cardíaca foram coletadas no instituto do coração de São Paulo (INCOR – SP) e no hospital metropolitano de João Pessoa. Por outro lado, as vozes dos outros 58 indivíduos saudáveis foram coletadas em ambiente extra-hospitalar. Ademais, o dispositivo utilizado para a gravação das vozes foi o gravador de áudio digital PX440, produzido pela Sony. Para análise dos dados coletados, foram aplicadas as seguintes técnicas para extrair e selecionar as características dos sinais: análise estatística, transformada rápida de Fourier, transformada discreta de wavelet e análise mel-cepstral. Por meio dessas técnicas, foram selecionadas características que serviram para alimentar as redes neurais artificiais (RNAs) desenvolvidas para cada sexo. Por meio das arquiteturas estabelecidas das redes, foi alcançada uma eficiência global de 96,7 % para as duas RNAs. De modo a garantir a usabilidade do sistema criado, foi desenvolvido um aplicativo de computador. Utilizando esse software, foram obtidos os valores de 91,86%; 88,1% e 92,1% no cálculo da acurácia, sensibilidade e especificidade, respectivamente. Portanto, o sistema de identificação de insuficiência cardíaca apresentou resultados promissores e que merecem ser estudados mais a fundo a fim de aprimorar a ferramenta desenvolvida.

Palavras Chaves: Insuficiência cardíaca, diagnóstico, vozes, análise de sinais, redes neurais artificiais.

ABSTRACT

Heart failure is a disease that disables the heart from properly pumping blood to nourish the entire body. Currently, the main diagnostic methods for this pathology are performed clinically through the measurement of B-type natriuretic peptide (BNP). As cardiovascular diseases are the main causes of premature death, the development of new technologies to identify these diseases is of great importance. Thus, this research presents the development of an identification system for the vocal distortions caused by heart failure in an individual. For the development of the software, the voices of 142 individuals were collected, separated by sex and age. Among these 142, 84 voices of people already diagnosed with heart failure were collected at the Heart Institute of Sao Paulo University (INCOR – USP) and at the Metropolitan Hospital of Joao Pessoa. On the other hand, the voices of the other 58 healthy individuals were collected in an extra-hospital environment. Furthermore, the device used for recording the voices was the PX440 digital audio recorder, produced by Sony. To analyze the collected data, the following techniques were applied to extract and select the characteristics of the signals: statistical analysis, fast Fourier transform, discrete wavelet transform and mel-cepstral analysis. By using these techniques, it was selected features to feed the artificial neural networks (ANNs) developed for each sex. Using the established architectures of the networks, an overall efficiency of 96.7% was achieved for both ANNs. In order to guarantee the usability of the created system, a computer application was developed. Using this software, values of 91.86%; 88.1% and 92.1% were obtained in the calculation of accuracy, sensitivity and specificity, respectively. Therefore, the heart failure identification system showed promising results that need to be further studied in order to improve the developed tool.

Key Words: Heart failure, diagnosis, voices, signal analysis, artificial neural networks.

SUMÁRIO

CAPÍTULO I – INTRODUÇÃO	15
1.1 INTRODUÇÃO	15
1.2 OBJETIVO GERAL	17
1.3 OBJETIVOS ESPECÍFICOS	17
1.4 ORGANIZAÇÃO DA DISSERTAÇÃO	17
1.6 MOTIVAÇÃO	Erro! Indicador não definido.
1.7 CONCLUSÃO	18
CAPÍTULO II – ESTADO DA ARTE	19
2.1 DIAGNÓSTICO DE INSUFICIÊNCIA CARDÍACA	19
2.1.1 Classificação de Acordo com a Fração de Ejeção	19
2.1.2 Classificação de Acordo com a Gravidade dos Sintomas	20
2.2 UTILIZAÇÃO DA RNA PARA DIAGNOSTICAR IC	21
2.3. IDENTIFICAÇÃO DA IC UTILIZANDO ANÁLISE DA VOZ	22
2.4. CONCLUSÃO	24
CAPÍTULO III – REFERENCIAL TEÓRICO	25
3.1 SISTEMA CARDIOVASCULAR	25
3.1.1 Aspectos Gerais do Coração	25
3.1.1.1 Funcionamento e Anatomia do Coração	26
3.1.1.2 Atividade Elétrica do Coração	28
3.1.1.3 Insuficiência Cardíaca	29
3.2 EXAMES	30
3.3 ANÁLISE DE SINAIS	31
3.3.1 Transformada Rápida de Fourier (FFT)	32
3.3.2 Transformada de Fourier de Tempo Curto (STFT)	32

3.3.3 Transformada Wavelet	33
3.3.4 Transformada Wavelet Contínua (CWT)	34
3.3.5 Transformada Wavelet Discreta (DWT)	35
3.3.5.1 As Famílias Wavelet	38
3.3.6 Análise Mel-Cepstral	40
3.4 REDES NEURAIS	43
3.4.1 Modelo de um Neurônio Artificial	44
3.4.2 Funções de Ativação	45
3.4.2.1 Função Linear	45
3.4.2.2 Função Limiar	46
3.4.2.3 Função Sigmóide	47
3.4.2.4 Função Softmax	49
3.4.3. Redes de Camada Simples	50
3.4.5 Redes de Múltiplas Camadas	50
3.4.6 Processo de Aprendizado Da Rede	51
3.5 CONCLUSÃO	52
CAPÍTULO IV – MATERIAIS E MÉTODOS	53
4.1 FLUXOGRAMA DA PESQUISA	53
4.2 AQUISIÇÃO DE DADOS	54
4.3 CRITÉRIOS PARA PARTICIPAÇÃO NA PESQUISA	56
4.4 PREPARAÇÃO DOS DADOS	57
4.5 PROCESSAMENTO DOS DADOS	60
4.5.1 Transformada Wavelet Discreta (DWT)	61
4.5.2 Dados Estatísticos	65
4.5.3 Transformada Rápida de Fourier (FFT)	66
4.5.4 Análise Mel-Cepstral	67
4.6 DESENVOLVIMENTO DA RNA	68

4.7 CONCLUSÃO	69
CAPÍTULO V – RESULTADOS	71
5.1 EXTRAÇÃO E SELEÇÃO DAS CARACTERÍSTICAS DAS VOZES MASCULINAS	71
5.1.1 Análise Multirresolução dos Sinais	71
5.1.2 Análise Estatística	76
5.1.3 Análise no Domínio da Frequência	83
5.1.4 Análise Mel-Cepstral	88
5.2 REDE NEURAL PARA OS INDIVÍDUOS DO SEXO MASCULINO	94
5.3 EXTRAÇÃO E SELEÇÃO DAS CARACTERÍSTICAS DAS VOZES FEMININAS	96
5.3.1 Análise Multirresolução dos Sinais	96
5.3.2 Análise Estatística	100
5.3.3 Análise no Domínio da Frequência	106
5.3.4 Análise Mel-Cepstral	112
5.4 REDE NEURAL PARA OS INDIVÍDUOS DO SEXO FEMININO	118
5.5 DESENVOLVIMENTO E FUNCIONAMENTO DO APLICATIVO	120
5.6 ACURÁCIA, SENSIBILIDADE E ESPECIFICIDADE DO SISTEMA	122
5.7 CONCLUSÃO	124
CAPÍTULO VI – CONCLUSÕES	125
6.1 CONCLUSÕES	125
6.2 SUGESTÕES PARA TRABALHOS FUTUROS	126
REFERÊNCIAS BIBLIOGRÁFICAS	128

ÍNDICE DE FIGURAS

Capítulo III

Figura 3. 1 - Fluxo Sanguíneo Através do Bombeamento Realizado pelo Coração	26
Figura 3. 2 - Representação das Camadas do Coração	27
Figura 3. 3 - Valvas do coração	27
Figura 3. 4 - Sistema de Gênese e Condução da Excitação Cardíaca	29
Figura 3. 5 - Representação de um sistema com uma entrada e uma saída	31
Figura 3. 6 - Análise Comparativa das Características de Escala da STFT e da Wavelet	34
Figura 3. 7 - Representação da Análise Wavelet em Multirresolução	36
Figura 3. 8 - Decomposição de um Sinal em Três Níveis	37
Figura 3. 9 - Wavelet de Haar.....	38
Figura 3. 10 - Representação das funções escala (esquerda) e wavelet-mãe (direita) da família Daubechies.....	39
Figura 3. 11 - Funções escala e wavelet da família Coiflet de ordem 1.....	40
Figura 3. 12 - Funções Escala e Wavelet da família Symmlet	40
Figura 3. 13 - Etapas de obtenção dos MFCCs.	43
Figura 3. 14 - Neurônio Artificial	44
Figura 3. 15 - Representação Gráfica da Função Linear.....	46
Figura 3. 16 - Representação Gráfica da Função Logística.....	48
Figura 3. 17 - Representação Gráfica da Função Tangencial Hiperbólica.	49
Figura 3. 18 - Rede Neural Artificial de Camada Simples.....	50
Figura 3. 19 - Rede Neural Artificial de Múltiplas Camadas.	51

Capítulo IV

Figura 4. 1 - Fluxograma da Pesquisa.....	54
Figura 4. 2 - Gravador digital PX440 da Sony.....	56
Figura 4. 3 - Distribuição das classes por idade – Masculino (IC).....	58
Figura 4. 4 - Distribuição das classes de idade – Masculino (Saudável).....	58
Figura 4. 5 - Distribuição das classes de idade – Feminino (IC).....	59
Figura 4. 6 - Distribuição das classes de idade – Feminino (Saudável).....	59
Figura 4. 7 – Análise em multirresolução do sinal sonoro de uma voz feminina.	62
Figura 4. 8 - Detalhe 1 do sinal sonoro da voz.....	62

Figura 4. 9 - Detalhe 2 do sinal sonoro da voz.	63
Figura 4. 10 - Detalhe 3 do sinal sonoro da voz.	64
Figura 4. 11 - Detalhe 4 do sinal sonoro da voz.	64

Capítulo V

Figura 5.1 - Evolução energética dos detalhes – Masculino.....	72
Figura 5.2 - Seleção dos detalhes – Masculino.	73
Figura 5.3 - Detalhe 4 – Masculino.	74
Figura 5.4 - Detalhe 5 – Masculino.	74
Figura 5.5 - Detalhe 6 – Masculino.	75
Figura 5.6 - Detalhe 7 – Masculino.	76
Figura 5.7 - Média dos detalhes – Masculino.	77
Figura 5.8 - Mediana dos detalhes – Masculino.	78
Figura 5.9 - Variância dos detalhes – Masculino.....	78
Figura 5.10 - Desvio Padrão dos detalhes – Masculino.....	79
Figura 5.11 - RMS dos detalhes – Masculino.	80
Figura 5. 12 - Amplitude dos detalhes – Masculino.....	80
Figura 5. 13 - Coeficiente de Pearson dos detalhes – Masculino.....	81
Figura 5.14 - Coeficiente de Variação dos detalhes – Masculino.	82
Figura 5.15 - FFT do detalhe 4 – Masculino.....	83
Figura 5. 16 - FFT do detalhe 5 – Masculino.....	84
Figura 5.17 - FFT do detalhe 6– Masculino.....	85
Figura 5.18 - FFT do detalhe 7 – Masculino.....	86
Figura 5.19 - Energia dos detalhes – Masculino.....	87
Figura 5.20 - Amplitude dos MFCCs – Masculino.....	89
Figura 5. 21 - Coeficiente de Pearson dos MFCCs – Masculino.	90
Figura 5.22 - Coeficiente de variação dos MFCCs – Masculino.	90
Figura 5.23 - Desvio padrão dos MFCCs – Masculino.	91
Figura 5.24 - Média dos MFCCs – Masculino.	92
Figura 5.25 - Mediana dos MFCCs – Masculino.....	92
Figura 5.26 - RMS dos MFCCs – Masculino.....	93
Figura 5. 27 - Variância dos MFCCs – Masculino.	93
Figura 5.28 - Rede Neural Artificial Desenvolvida – Masculino.....	95
Figura 5.29 - Evolução energética dos detalhes – Feminino.	97

Figura 5.30 - Seleção dos detalhes – Feminino.	97
Figura 5.31 - Detalhe 4 – Feminino.	98
Figura 5.32 - Detalhe 5 – Feminino.	99
Figura 5.33 - Detalhe 6 – Feminino.	99
Figura 5.34 - Detalhe 7 – Feminino.	100
Figura 5.35 - Amplitude dos detalhes – Feminino.	101
Figura 5. 36 - Coeficiente de Pearson dos detalhes – Feminino.	101
Figura 5. 37 - Coeficiente de variação dos detalhes – Feminino.....	102
Figura 5. 38 - Desvio padrão dos detalhes – Feminino.....	103
Figura 5.39 - Média dos detalhes – Feminino.	103
Figura 5. 40 - Mediana dos detalhes – Feminino.....	104
Figura 5. 41 - Valor RMS dos detalhes – Feminino.....	105
Figura 5. 42 - Variância dos detalhes – Feminino.	105
Figura 5. 43 - FFT do detalhe 4 – Feminino.	107
Figura 5. 44 - FFT do detalhe 5 – Feminino.	108
Figura 5.45 - FFT do detalhe 6 – Feminino.	109
Figura 5.46 - FFT do detalhe 7 – Feminino.	110
Figura 5. 47 - Energia dos detalhes – Feminino.	111
Figura 5.48 - Amplitude dos MFCCs – Feminino.	112
Figura 5. 49 - Segundo coeficiente de Pearson dos MFCCs – Feminino.	113
Figura 5. 50 - Coeficiente de Variação dos MFCCs – Feminino.	114
Figura 5.51 - Desvio Padrão dos MFCCs – Feminino.....	115
Figura 5.52 - Média dos MFCCs – Feminino.	115
Figura 5.53 - Mediana dos MFCCs – Feminino.	116
Figura 5.54 - Valor RMS dos MFCCs – Feminino.	117
Figura 5.55 - Variância dos MFCCs – Feminino.....	117
Figura 5. 56 - Rede Neural Artificial Desenvolvida – Feminino.	119
Figura 5. 57 - Aplicativo Desenvolvido para Análise das Vozes – Feminino.....	120
Figura 5. 58 - Classificação Gerada pelo Aplicativo – Feminino.	121

ÍNDICE DE TABELAS

Capítulo II

Tabela 2.1 - Classificação da IC com ICFeR, ICFei e ICFeP	20
Tabela 2. 2 - Classificação Segundo a New York Heart Association	21
Tabela 2.3 - Acurácias médias obtidas utilizando diferentes classificadores.....	24

Capítulo III

Tabela 3. 1 - Classe Funcional de acordo com a American Heart Association	30
---	----

Capítulo IV

Tabela 4. 1 – Especificações técnicas do gravador digital PX440 da Sony.....	55
Tabela 4. 2 - Critérios de Exclusão e Inclusão.	57
Tabela 4. 3 – Análise de classe.	60
Tabela 4. 4 - Faixa de Frequência dos Detalhes.	61

Capítulo V

Tabela 5. 1 - Características relevantes para o desenvolvimento da RNA – Masculino.....	83
Tabela 5. 2 - Análise no domínio da frequência – Masculino.	87
Tabela 5. 3 - Análise dos coeficientes cepstrais – Masculino.....	94
Tabela 5. 4 - Performance da RNA para diferentes quantidades de neurônios na camada oculta – Masculino.	95
Tabela 5. 5 - Características relevantes para o desenvolvimento da RNA – Feminino.	106
Tabela 5. 6 - Análise no domínio da frequência – Feminino.....	111
Tabela 5. 7 - Análise dos coeficientes cepstrais – Masculino.....	118
Tabela 5. 8 - Performance da RNA para diferentes quantidades de neurônios na camada oculta – Feminino.....	119
Tabela 5. 9 - Resultados Obtidos dos Testes – Feminino.	122
Tabela 5. 10 - Acurácia, Sensibilidade e Especificidade do sistema – Feminino.	123
Tabela 5. 11 - Análise Comparativa das técnicas empregadas para o diagnóstico de IC – Feminino.	123

CAPÍTULO I

INTRODUÇÃO

Neste capítulo, será apresentada a contextualização do tema proposto da pesquisa, assim como sua importância para a área de estudo, com o intuito de demonstrar a contribuição da presente pesquisa. Além disso, são expostos os objetivos, tanto gerais quanto específicos, e a maneira como a presente dissertação foi organizada.

1.1 INTRODUÇÃO

A Organização Pan-Americana da Saúde (OPAS) afirma que, em 2016, cerca de 17,9 milhões de pessoas morreram devido as doenças cardiovasculares (DCVs) no mundo; o que corresponde a 31% das mortes daquele ano. Nesse sentido, tais dados colocam as DCVs como as principais causas de mortes prematuras (abaixo dos 70 anos), ultrapassando câncer e guerras. Além disso, apesar das DCVs serem um problema global, os países mais pobres são os mais afetados por esse problema. No Brasil, de acordo com a secretaria de saúde do estado de Goiás, ocorreram 169.557 mortes por doenças cardiovasculares no Brasil.

Ao analisar especificamente a insuficiência cardíaca (IC), a agência americana de controle e prevenção de doenças (CDC) estima que aproximadamente 6,2 milhões de americanos tenham IC, o que causou 379.800 mortes em 2018 e custou aos cofres públicos e privados 30,7 bilhões de dólares em 2012. No Brasil, a IC correspondeu, em 2015, a 5,51% das causas de internação hospitalar em pessoas acima de 60 anos, sendo a segunda maior causa de internação nessa faixa etária, atrás apenas de pneumonia (ROSSETO et. al., 2019).

Apesar da melhoria nos resultados com terapia médica, as taxas de admissão e readmissão por IC permanecem altas. Como a redução dessas taxas podem simultaneamente reduzir custos e melhorar a qualidade do atendimento, sistemas de saúde públicos e privados têm cada vez mais o tratamento preventivo e o diagnóstico precoce como foco de pesquisa e investimento (DESAY e STEVENSON, 2012).

No que concerne ao diagnóstico da IC, a Sociedade Brasileira de Cardiologia (SBC) apresenta as seguintes diretrizes:

- 1- Avaliação clínica: Consiste em identificar os sintomas de insuficiência cardíaca (falta de ar, fadiga, ganho de peso e inchaço dos membros inferiores e superiores decorrente do acúmulo de líquido no organismo).
- 2- Ecocardiograma: Permite a avaliação da função ventricular sistólica (ejeção de sangue) esquerda e direita, da função diastólica (regula o enchimento ventricular), das espessuras parietais, do tamanho das cavidades, da função valvar, da estimativa hemodinâmica não invasiva e das doenças do pericárdio.
- 3- Peptídeos natriuréticos: Avalia os biomarcadores estudados em IC, destacam-se os peptídeos natriuréticos (BNP e NT-proBNP).

Apesar da comprovada eficácia das diretrizes apresentadas, a identificação da IC se dá por meio da avaliação conjunta desses exames; o que pode dificultar o diagnóstico, visto que doenças como diabetes, insuficiência hepática e pneumopatia podem acarretar em sintomas semelhantes.

Assim, os estudos que objetivam o avanço do reconhecimento e do tratamento da IC apresentam grande relevância. Dessa forma, diversas pesquisas têm se direcionado para o desenvolvimento de sistemas inteligentes que visem auxiliar no diagnóstico preciso dessa doença.

Embora tenha havido notório avanço tecnológico no diagnóstico da insuficiência cardíaca nas últimas décadas, algumas áreas e possibilidades do conhecimento ainda não foram devidamente abordadas. Nesse sentido, o estudo da voz humana e sua mudança com o aparecimento de algum mau funcionamento do sistema cardiovascular merece uma atenção especial por sua fácil implementação.

De acordo com o estudo realizado por Murton et al. (2017), a alteração da voz de uma pessoa pode sinalizar uma mudança fisiológica do paciente e apontar o aparecimento de uma possível doença cardíaca.

Portanto, este estudo propõe uma análise detalhada da voz de pessoas com insuficiência cardíaca, utilizando as seguintes técnicas de tratamento de sinais: análise estatística da voz, transformada rápida de Fourier (FFT), transformada de Wavelet Discreta (DWT) e análise mel-cepstral do sinal. As características do sinal obtidas através dessas técnicas serão utilizadas no processo de desenvolvimento de uma rede neural artificial (RNA) capaz de identificar, de forma automática, as alterações na voz causadas pela IC.

1.2 OBJETIVO GERAL

Desenvolver um sistema capaz de auxiliar a identificação da insuficiência cardíaca por meio das alterações nas vozes dos indivíduos com a utilização de RNAs.

1.3 OBJETIVOS ESPECÍFICOS

A fim de atingir o objetivo geral, as seguintes etapas foram estabelecidas:

- Analisar a relação entre insuficiência cardíaca e alterações nos padrões da voz;
- Desenvolver uma RNA capaz de identificar tais alterações;
- Desenvolver um aplicativo de computador para auxiliar o profissional de saúde na identificação de IC.
- Convalidar o sistema desenvolvido.

1.4 ORGANIZAÇÃO DA DISSERTAÇÃO

O presente trabalho é constituído de 6 capítulos, organizados da seguinte maneira:

Capítulo I: Contextualiza o trabalho, por meio da introdução, esclarece seus objetivos, gerais e específicos, e o modo como a dissertação foi organizada.

Capítulo II: Apresenta a contextualização do tema proposto da pesquisa, assim como sua importância e trabalhos realizados na área.

Capítulo III: Aborda o referencial teórico necessário para a elaboração deste trabalho. Engloba conteúdos como o sistema cardíaco e os exames realizados para o diagnóstico da IC atualmente. Além de assuntos relativos ao processamento de sinais sonoros.

Capítulo IV: Apresenta a metodologia adotada, descrevendo os procedimentos utilizados, desde a obtenção dos sinais, passando pela elaboração das RNAs, até o desenvolvimento final do aplicativo.

Capítulo V: Faz um levantamento dos resultados obtidos, assim como uma discussão sobre os mesmos.

Capítulo VI: Fornece as conclusões do trabalho e sugestões para projetos futuros.

1.6 CONCLUSÃO

Nesse capítulo foi realizada uma introdução do presente trabalho, de seus objetivos (geral e específicos), de sua organização e da motivação da realização da pesquisa.

Na introdução, foi realizada uma explanação geral sobre a atual situação global e brasileira da insuficiência cardíaca. Além disso, foi exposta as diretrizes de diagnóstico proferida pela SCB. Por fim, apresentou-se a proposta da presente pesquisa que concerne em uma técnica inovadora de auxílio ao diagnóstico da IC.

Na seção dos objetivos, foram expostos os objetivos geral e específicos desta dissertação. Por outro lado, na organização da dissertação, foi abordada a forma como os capítulos estavam alocados ao longo do trabalho.

Por fim, apresenta-se a justificativa desta pesquisa, na qual as motivações para a realização deste trabalho são expostas. Tais motivações incluem: auxiliar o médico na identificação da IC e diminuir a mortalidade da IC por meio do seu diagnóstico precoce.

CAPÍTULO II

ESTADO DA ARTE

Neste capítulo é apresentado o desenvolvimento do tema ao longo do tempo por diversos autores. Aqui, visa-se mostrar o atual estado de aplicação das tecnologias de auxílio do diagnóstico da insuficiência cardíaca com o objetivo expressar a contribuição deste estudo.

2.1 DIAGNÓSTICO DE INSUFICIÊNCIA CARDÍACA

De acordo com a Sociedade Europeia de Cardiologia, a insuficiência cardíaca (IC) é uma síndrome caracterizada pela presença de alguns sintomas típicos – falta de ar, inchaço de tornozelos e fadiga, por exemplo – que podem ser acompanhados por elevada pressão venosa jugular, estertores pulmonares e edema periférico. Em geral, a IC é causada por uma anormalidade estrutural cardíaca, incapacitando o coração de atender às necessidades metabólicas tissulares, ou poder fazê-lo por meio de elevadas pressões intracardíacas (PONIKOWSKI et al., 2016).

O Departamento de Insuficiência Cardíaca (DEIC) da Sociedade Brasileira de Cardiologia (SBC) apresentou em 2018 uma versão atualizada da Diretriz de Insuficiência Cardíaca Crônica, objetivando a diminuição da taxa de mortalidade associada à IC no Brasil. Para tanto, o referido documento apresenta uma série de procedimentos para o diagnóstico e tratamento precoce de pacientes em risco, considerando duas formas de classificação da IC: classificação de acordo com a fração de ejeção e classificação de acordo com a gravidade dos sintomas.

2.1.1 Classificação de Acordo com a Fração de Ejeção

De acordo com a SBC (2018), a terminologia mais tradicional para a classificação da IC é a que se baseia na fração de ejeção do ventrículo esquerdo (FEVE), onde o volume de sangue que

deixa o coração a cada batimento é calculado em termos percentuais. Nesse sentido, uma FEVE maior ou igual a 50% representa uma fração de ejeção preservada (ICFEp); uma FEVE menor que 40% representa uma fração de ejeção reduzida (ICFEr) e uma FEVE que apresenta um valor entre 40 e 49% é classificada como fração de ejeção intermediária (ICFEi).

Em termos de diagnóstico e classificação, a tabela 2.1 apresenta os critérios estabelecidos pela Sociedade Europeia de Cardiologia a serem observados pelos profissionais da saúde.

Classificação da IC	ICFEr	ICFEi	ICFEp
Critérios	1	Sintomas	Sintomas
	2	FEVE < 40%	FEVE ≥ 50
	3	–	–
		1. Elevado peptídeo natriurético do tipo B (BNP);	1. Elevado peptídeo natriurético do tipo B (BNP);
		2. Pelo menos um critério adicional	2. Pelo menos um critério adicional
		a) Doença estrutural do coração	a) Doença estrutural do coração
		b) Disfunção diastólica	b) Disfunção diastólica

Tabela 2.1 - Classificação da IC com ICFEr, ICFEi e ICFEp. Fonte: PONIKOWSKI et al. (Adaptado), 2016.

Assim, de acordo com a tabela 2.1, um paciente sintomático que apresente ICFEr é classificado como um indivíduo portador de IC. Por outro lado, um paciente que apresente sintomas associados a um ICFEi ou ICFEp deve apresentar elevado BNP e mais um critério adicional (Doença cardíaca estrutural ou disfunção diastólica) para que seja classificado como portador de IC.

2.1.2 Classificação de Acordo com a Gravidade dos Sintomas

Quando se considera a gravidade dos sintomas para a classificação da IC, o sistema de classes funcionais desenvolvido pela *New York Heart Association* (NYHA) é um dos métodos mais aplicados da atualidade. Nesse sistema, a classe funcional (CF) NYHA I corresponde a pacientes que não apresentam sintomas ao desempenhar atividades físicas do dia-a-dia; a CF NYHA II se refere aos pacientes com sintomas leves ao realizar atividades físicas do cotidiano; pacientes pertencentes à CF NYHA III manifestam sintomas moderados quando desenvolvem atividades diárias, porém apresentam conforto quando estão em descanso; por último, a CF NYHA

IV inclui os pacientes incapazes de realizar qualquer atividade física sem apresentar os sintomas da IC; ademais, esses pacientes podem apresentar sintomas quando estão em descanso (GIUSTINO et al., 2020). Com o intuito de facilitar a classificação segundo a *New York Heart Association*, a tabela 2.2 apresenta um resumo desse sistema.

Classificação	Descrição
NYHA I	Pacientes sem limitações para realizar atividades físicas
NYHA II	Pacientes com poucas limitações para realizar atividades físicas com a presença de sintomas (fadiga, dispneia e palpitação).
NYHA III	Pacientes com sintomas moderados (fadiga, dispneia e palpitação) para realizar atividades físicas ordinárias, mas não apresentam sintomas quando estão em repouso.
NYHA IV	Pacientes incapazes de realizar atividades físicas sem apresentar sintomas de IC. Além disso, esses pacientes apresentam sintomas quando estão em repouso.

Tabela 2. 2 - Classificação Segundo a New York Heart Association. Fonte: RESNIK; MISCH (Adaptado), 2017.

Dessa forma, percebe-se que a classificação de acordo com a classe funcional do paciente é determinada via análise clínica do médico para com seu paciente. Portanto, tal método está sujeito a erro humano tanto na descrição dos sintomas pelo paciente quanto no diagnóstico gerado pelo médico.

2.2 UTILIZAÇÃO DA RNA PARA DIAGNOSTICAR IC

Apesar dos métodos tradicionais de identificação e classificação da IC apresentarem bons resultados, esses processos são, em sua maioria, de alto custo e suscetíveis ao erro humano. Dessa forma, o desenvolvimento de sistemas inteligentes que auxiliem o médico no diagnóstico precoce da IC é uma demanda da medicina mundial.

Nesse sentido, com foco no diagnóstico rápido e eficiente, Van Gils et al. (1997) aplicaram redes neurais artificiais (RNA) para determinar variáveis específicas que ajudem a definir o estado clínico dos pacientes. Os parâmetros utilizados na pesquisa de Van Gils foram: pressão arterial; pressão arterial pulmonar; frequência cardíaca; entre outros. Apesar desse estudo determinar uma combinação de dados que pode levar a um avanço no diagnóstico de insuficiência cardíaca, ela não fornece uma combinação de dados para tomada de decisão.

Ainda aplicando RNA, Acharya et al. (2008) realizaram uma análise espectral de um sinal obtido via eletrocardiograma (ECG). Nessa pesquisa, a transformada rápida de Fourier (FFT) foi utilizada em conjunto com modelos estatísticos, tais como: Média Móvel (MA), processos Auto Regressivos (AR) e processos de Médias Móveis Auto Regressivos (ARMA). O estudo concluiu que a FFT não é a ferramenta mais adequada para a análise da frequência cardíaca, devido a característica de não estacionariedade dos sinais obtidos pelo ECG. Além disso, constatou-se que o modelo estatístico ARMA apresenta a melhor entrada para a RNA utilizada na pesquisa, obtendo uma acurácia de 83,83%.

Por fim, em 2014; Karaye, Saminu e Ozkurt realizaram uma análise espectral de sinais de ECGs de pessoas saudáveis e de pessoas com problemas de insuficiência cardíaca usando a teoria dos sistemas não lineares como variância, curtose e assimetria. As doenças que causam a insuficiência cardíaca foram identificadas a uma taxa de 94,9% de acurácia por uma RNA com apenas uma camada oculta com 15 neurônios.

Assim, com base no exposto, é possível perceber que a utilização de sistemas inteligentes que dão suporte ao profissional de saúde na avaliação do paciente é uma abordagem promissora na busca de um melhor gerenciamento dos sintomas no estágio inicial da doença, visto que a maioria dos casos de insuficiência cardíaca aguda tem como principais fatores a falta de cuidados pessoais e a falha da gestão dos sintomas (LEE et al., 2017).

2.3. IDENTIFICAÇÃO DA IC UTILIZANDO ANÁLISE DA VOZ

Uma das principais características da insuficiência cardíaca é a presença de edemas em várias regiões do corpo, principalmente nos membros inferiores e, nos casos mais graves, nos pulmões. Como sintoma de sua presença, os edemas causam retenção excessiva de líquido nos tecidos do corpo, acarretando em inchaço e ganho de peso do paciente (PONIKOWSKI et al., 2016). Além disso, a presença de edemas nos pulmões causa uma modificação estrutural em todo o sistema respiratório, causando, por consequência, alteração na fala do indivíduo.

Nesse sentido, em estudo realizado em 2017, Murton et al. conseguiram correlacionar a presença da insuficiência cardíaca com a mudança de alguns parâmetros da voz dos pacientes. O estudo apresentou os seguintes resultados objetivos:

- 1- A qualidade vocal na maioria dos pacientes aumentou após o tratamento da doença na leitura de uma passagem específica de texto;

Essa característica se refere à capacidade de gerar voz limpa, isto é, sem chiado excessivo. Nesse sentido, a presença do chiado pode ser associada a patologias – laringomalácia e Estenose subglótica, por exemplo – que restringem a passagem de ar pelas cordas vocais.

- 2- As vozes dos pacientes tratados tenderam a ser mais uniformes;

Essa característica foi observada devido a uma diminuição dos valores de desvio padrão da frequência fundamental da voz da maioria dos pacientes após o tratamento da insuficiência cardíaca. Além disso, houve um decréscimo do desvio da amplitude dos sinais sonoros, o que contribui para a uniformidade da voz.

- 3- A maioria dos pacientes precisaram de um tempo menor para ler a passagem após o tratamento e produziu mais fonemas por frase.

Em geral, os pacientes tratados tinham a capacidade de falar mais rápido ou precisavam de uma menor quantidade de paradas para respirar ao ler uma passagem de texto.

Dessa forma, percebe-se que o trabalho desenvolvido por Murton mostra resultados promissores, pois apresenta características interessantes para uma possível identificação da IC. Porém, já outros parâmetros (como a energia associada a uma determinada frequência e análise estatística dos dados) da voz podem ser identificados com técnicas de processamento de sinais.

Em outro estudo, Reddy et. al. (2021) apresentou um sistema de identificação de IC por meio da análise de sinais de fala. Para o desenvolvimento dessa pesquisa, os autores analisaram as vozes de 25 indivíduos saudáveis e 20 pacientes com IC. Além disso, o método utilizado pelos autores foi por meio da aplicação combinada dos coeficientes cepstrais (MFCCs) e das características extraídas dos sinais de fonte glótica. Ademais, a aquisição de dados se deu por meio da leitura repetida por 3 vezes de um texto de 91 palavras, o que levou em média 1 minuto para cada gravação.

Assim, a pesquisa realizada por Reddy apresentou diferentes resultados para diferentes classificadores, como é apresentado na tabela 2.3.

Classificador	Acurácia (%)
---------------	--------------

Máquinas de vetores de suporte	77,15
<i>Extra trees</i>	72,29
<i>Data Boost</i>	79,23
Rede neural artificial	81,51

Tabela 2.3 - Acurácias médias obtidas utilizando diferentes classificadores. Fonte: REDDY et. al., 2021 (Adaptado).

Portanto, nota-se que a utilização da análise sonora da voz para auxiliar o diagnóstico precoce da insuficiência cardíaca é uma realidade do cenário atual. No entanto, sua aplicação ainda tem uma ampla área a ser explorada, seja por diferentes técnicas de análises de sinal, seja por diversas formas de aplicação dos classificadores. Nesse sentido, o presente trabalho se propõe a utilizar técnicas de análise de sinal no domínio do tempo (estatística e MFCC), no domínio da frequência (FFT) e no domínio tempo-frequência (*wavelet*) em sinais sonoros de vozes de pessoas saudáveis e de pessoas com IC para extrair características desses sinais e aplicar como entrada em uma RNA.

2.4. CONCLUSÃO

O capítulo 2 deste trabalho se refere à evolução do diagnóstico da IC. Nesse sentido, o DEIC da SBC apresenta a classificação da IC de acordo com a fração de ejeção e a classificação de acordo com a gravidade dos sintomas.

Com a evolução da tecnologia, alguns sistemas inteligentes foram desenvolvidos com o intuito de identificar a IC de forma rápida e precisa. No geral, esses sistemas utilizam sinais de ECGs, como mostrado nos trabalhos de Acharya et al. (2008) e de Karye (2014).

Por último, apresenta-se os estudos que relacionam voz e IC. Nesse sentido, os estudos de Murton et. al. (2017) e de Reddy et. al. (2021) demonstram importantes avanços nessa área de estudo.

Portanto, esse capítulo se refere ao desenvolvimento do diagnóstico da IC ao longo dos anos, passando por exames tradicionais e chegando aos estudos mais atuais que correlacionam IC com a mudança no padrão sonoro da voz do ser humano.

CAPÍTULO III

REFERENCIAL TEÓRICO

Neste capítulo serão abordados tópicos relacionados ao coração, com destaque para sua anatomia, uma vez que o seu possível mau funcionamento é objeto de interesse deste trabalho. Além disso, também serão apresentados conteúdos relacionados à análise de sinais e redes neurais artificiais de modo a cobrir toda a teoria necessária para o entendimento desta pesquisa.

3.1 SISTEMA CARDIOVASCULAR

O sistema cardiovascular tem como principal função transportar oxigênio, nutrientes, hormônios e enzimas entre as células do organismo. O oxigênio e os nutrientes possibilitam as células produzirem energia química, o que permite ao organismo um bom funcionamento. As enzimas, por sua vez, auxiliam as reações químicas no interior das células e os produtos residuais dessas reações são transportados até os órgãos como pulmões e rins (RIZZO, 2012). O sistema cardiovascular não possui contato com o meio externo e o seu órgão central.

3.1.1 Aspectos Gerais do Coração

Essa seção se dedica ao estudo da anatomia e funcionamento do coração. Assim, as principais funções e estruturas do coração serão abordadas nos próximos tópicos.

3.1.1.1 Funcionamento e Anatomia do Coração

O coração está posicionado obliquamente entre os pulmões, aproximadamente dois terços do seu volume encontram-se do lado esquerdo da linha média do corpo. Seu formato é de um cone fechado e seu tamanho é de um punho fechado (RIZZO, 2012).

Segundo Santos (2014), o coração realiza a função de fornecer a pressão adequada para que o sangue circule – através dos vasos sanguíneos – em todo o organismo, de modo a proporcionar nutrientes e oxigênio para todas as células do corpo. Em um ser humano adulto, um coração saudável bombeia aproximadamente 5 litros de sangue por minuto.

O sangue é bombeado de forma sequencial, onde o lado direito do coração força a circulação pulmonar e o seu lado esquerdo força a circulação sistêmica. A figura 3.1 ilustra a distribuição do sangue realizada pelo coração.

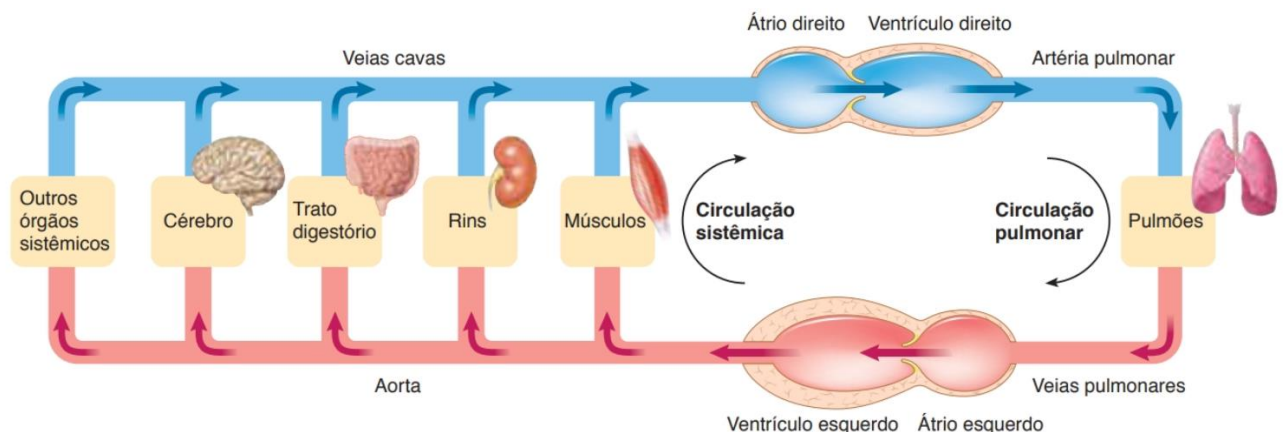


Figura 3. 1 - Fluxo Sanguíneo Através do Bombeamento Realizado pelo Coração. Fonte: Sherwood, 2011.

Ao realizar um corte no coração, é possível perceber que este é formado por camadas. A camada mais externa é constituída pelo saco pericárdio, composto por tecido conjuntivo e rígido que impede a hiperdistensão do coração e atua como membrana protetora. A camada intermediária é chamada de miocárdio e é basicamente uma camada muscular que tem como função proporcionar contrações musculares que ejetam o sangue das câmaras do coração. A última camada (a mais interna) é chamada de endocárdio e é caracterizada por ser uma camada fina de endotélio (tecido epitelial). O endocárdio tem como função atuar no revestimento interno protetor das câmaras e das valvas. A figura 3.2 mostra uma representação do coração e suas camadas (GRAAFF, 2002).

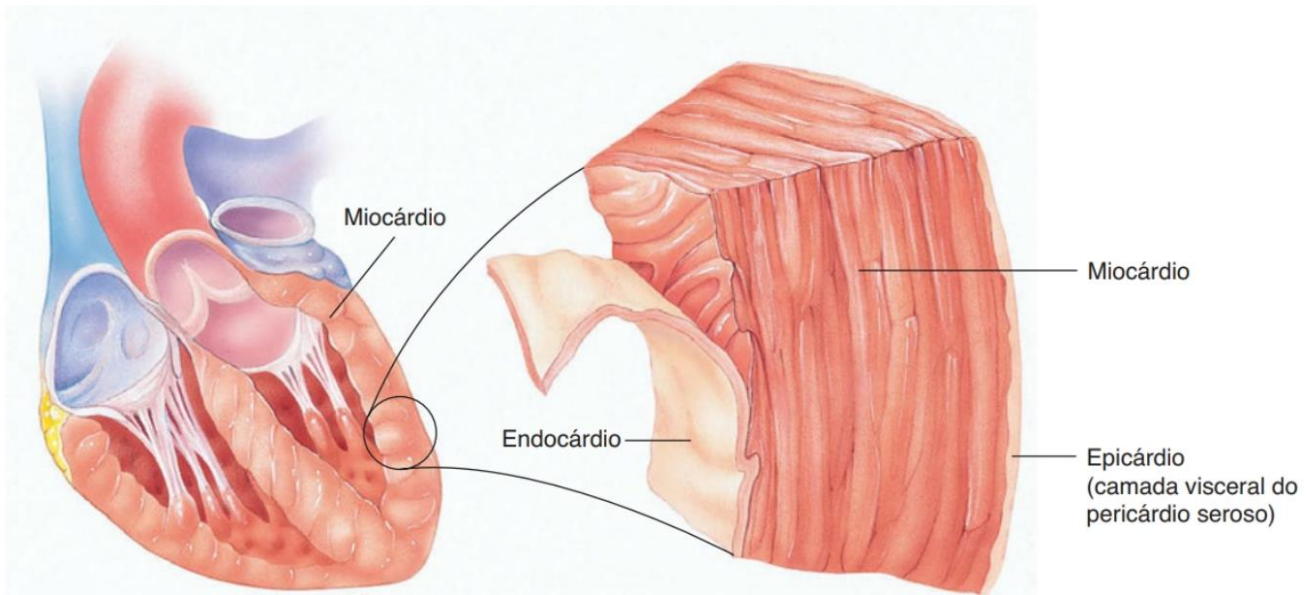


Figura 3. 2 - Representação das Camadas do Coração. Fonte: Rizzo, 2012

O interior do coração é composto por quatro câmaras que recebem o sangue das partes do corpo. As duas câmaras superiores recebem os nomes de átrio direito e átrio esquerdo. Por outro lado, as duas câmaras inferiores são chamadas de ventrículos direito e esquerdo. Os átrios são câmaras receptoras que recebem o sangue do corpo e dos pulmões. Os músculos cardíacos se contraem e ejetam o sangue dos átrios para os ventrículos. Para manter o fluxo do sangue em apenas uma direção, o coração disponibiliza de um mecanismo chamado de valvas (BECKER et. al., 2018). O sistema de valvas descrito é ilustrado nas figuras 3.3.

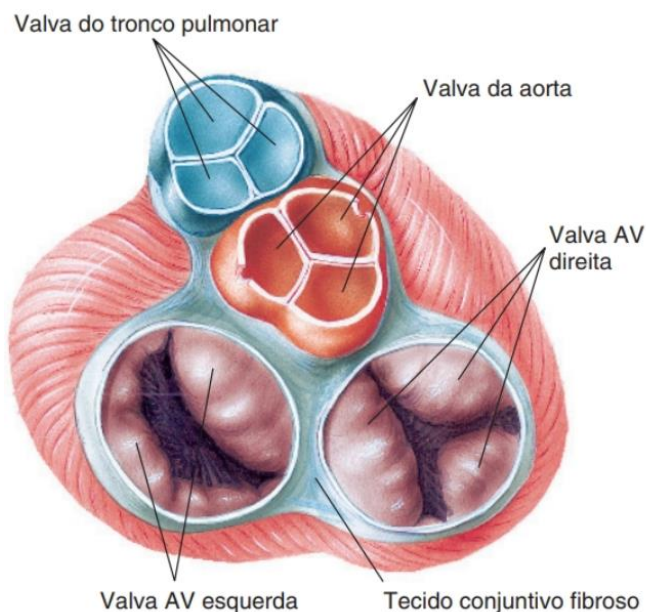


Figura 3. 3 - Valvas do coração. Fonte: GRAAFF, 2002.

Dentro desse sistema composto por átrios, ventrículos e valvas, o sangue circula da seguinte maneira: O átrio direito recebe o sangue venoso (pobre em oxigênio) da veia cava superior e o direciona para o ventrículo direito por meio da valva atrioventricular (AV) direita. Após a contração muscular do ventrículo, o sangue é ejetado para os pulmões, onde é oxigenado e transportado para o átrio esquerdo. Por meio da valva atrioventricular (AV) esquerda, o sangue é transportado para o ventrículo esquerdo e, após todo esse processo, o sangue oxigenado deixa o ventrículo esquerdo através da parte ascendente da aorta e é enviado para todo o organismo (GRAAFF, 2002).

3.1.1.2 Atividade Elétrica do Coração

Para a ejeção do sangue, o coração se contrai devido ao potencial de ação que percorre as membranas das células musculares. Assim, como resultado desse potencial de ação, ocorre o batimento rítmico do coração (SHERWOOD, 2011).

A atividade elétrica cardíaca tem como gatilho o nó sinoatrial, uma pequena região especializada na parede do átrio direito e perto da veia cava superior. Além disso, também no átrio direito, há o nó atrioventricular que retarda a condução do potencial de ação entre o miocárdio atrial e ventricular, esse nó se caracteriza por ser um pequeno feixe de células especializadas do músculo cardíaco. Partindo do nó atrioventricular, surge o tecido especializado em condução, chamado de Feixe de His, que se estende para a musculatura ventricular e forma uma extensa rede de condução intraventricular, as fibras de Purkinje (AIRES, 2018).

O sistema descrito pode ser observado na figura 3.4, onde à esquerda é mostrado o sistema que gera o potencial de ação e à direita é apresentado o sistema de propagação do estímulo.

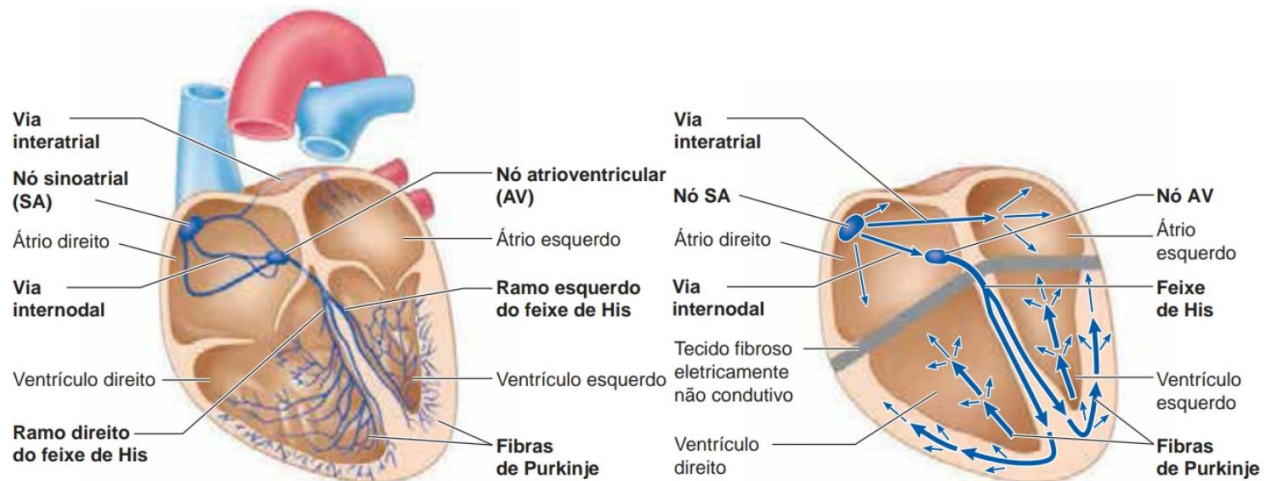


Figura 3. 4 - Sistema de Gênese e Condução da Excitação Cardíaca. Fonte: Sherwood, 2011.

Por fim, cabe destacar que o nó sinoatrial tem uma taxa de aproximada 70 a 80 potenciais de ação por minuto, induzindo os demais nós a funcionarem no mesmo ritmo, o que gera, em um coração saudável e em repouso, a uma taxa de 70 a 80 batimentos por minuto.

3.1.1.3 Insuficiência Cardíaca

A insuficiência cardíaca (IC) é uma síndrome clínica que surge em decorrência de alterações estruturais ou funcionais do coração, herdadas ou adquiridas ao longo da vida do indivíduo, que impossibilitam o suprimento das necessidades metabólicas do organismo ou que o faça acarretando o aumento das pressões de enchimento das suas câmaras (NOBRE, 2016). Ao ser acometido pela IC, o sujeito apresenta uma série de sintomas (dispneia e fadiga, por exemplo) e sinais (como edema e estertores) clínicos que aumenta a possibilidade de hospitalizações, piora na qualidade de vida e diminui a expectativa de vida (LOCALZO, 2014).

Para a classificação da IC, a *American Heart Association* introduziu, em 2005, um sistema de estadiamento para classificação de IC (CANNON e STEINBERG, 2012). Nessa metodologia, a determinação do estágio da IC pode ser feita de 4 formas diferentes, como exposto na tabela 3.1.

Capacidade Funcional	Avaliação Objetiva
----------------------	--------------------

Classe I	Pacientes com cardiopatia, mas sem limitações para atividades físicas. Atividades físicas normais não causam fadiga exagerada, palpitações, dispneia ou dor anginosa.
Classe II	Pacientes com cardiopatia que resulta em pequena limitação das atividades físicas. Não há desconforto em repouso. As atividades físicas normais causam fadiga, palpitação, dispneia ou dor anginosa
Classe III	Pacientes com cardiopatia que resulta em evidente limitação das atividades físicas. Não há desconforto em repouso. Atividades físicas mínimas causam fadiga, palpitação, dispneia ou dor anginosa
Classe IV	Pacientes com cardiopatia que os torna incapacitados para realizar qualquer atividade física sem desconforto. Os sintomas de insuficiência cardíaca ou de síndrome anginosa podem estar presentes mesmo durante o repouso. As tentativas de empreender qualquer atividade física resultam em aumento do desconforto

Tabela 3. 1 - Classe Funcional de acordo com a American Heart Association. Fonte: Loscalzo, 2014 (Adaptado).

Além da classificação segundo a *American Heart Association*, a IC também pode ser classificada a partir da fração de ejeção do ventrículo esquerdo (FEVE), medida pelo ecocardiograma em IC com FEVE reduzida ($< 40\%$), IC com FEVE preservada ($FEVE \geq 50\%$) e com FEVE intermediária ($40 - 49\%$) (NOBRE, 2016).

3.2 EXAMES

Primeiramente, o exame físico é indispensável para a determinação das causas da IC assim como para mensurar a gravidade da síndrome (LOSCALZO, 2014). Além disso, Cannon e Steinberg (2012) elencam alguns exames que são extensivamente utilizados no processo de identificação da IC; destacam-se os seguintes:

- a) **Exames Laboratoriais:** Em geral, utiliza-se o peptídeo natriurético do tipo B (BNP) na avaliação e na triagem de pacientes com dispneia aguda. Nesse sentido, estudos demonstram que o aumento dos níveis de BNP, em adição a outros parâmetros clínicos, podem confirmar o diagnóstico da IC como uma explicação para o achado de dispneia nos pacientes.

- b) Raio X de Tórax:** 50% dos pacientes com IC apresentam pressão capilar pulmonar aumentada. Além disso, derrames pleurais bilaterais e cardiomegalia também podem estar presentes nesses indivíduos.
- c) Ecocardiograma:** Ajuda na determinação do comprometimento ventricular esquerdo e/ou direito.
- d) Teste de Caminhada de 6 Minutos:** Uma menor distância atingida no teste pode está relacionadas a IC.

Por fim, destaca-se a crescente aplicação da análise de sinais no diagnóstico da IC. Nesse sentido, menciona-se o trabalho de Acharya et al. (2008) que utiliza o estudo de sinais obtidos de um eletrocardiograma para a identificação da IC. Portanto, o estudo das técnicas de análise de sinais é uma etapa importante para o desenvolvimento de ferramentas que auxiliem o profissional da saúde no diagnóstico da IC.

3.3 ANÁLISE DE SINAIS

O sinal é um conjunto de dados ou informações sobre um fenômeno ou acontecimento. Na maioria dos casos, os sinais são funções da variável independente tempo (LATHI, 2007).

Sinais são modificados por sistemas e, de acordo com Roberts (2009), quando um ou mais estímulos são aplicados a uma ou mais entradas do sistema, este produz uma ou mais respostas em suas saídas. A Figura 3.5 representa um sistema de uma entrada e uma única saída.

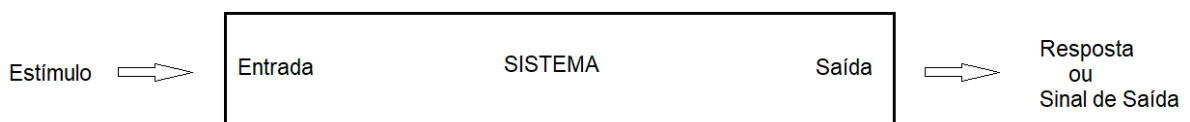


Figura 3. 5 - Representação de um sistema com uma entrada e uma saída. Fonte: Adaptado de ROBERTS, 2009.

A maioria dos dispositivos de aquisição de sinais fornecem valores em função do tempo, o que dificulta a interpretação do fenômeno que está querendo se obter informações. Sendo assim, algumas técnicas são aplicadas de forma a facilitar a análise desses sinais.

3.3.1 Transformada Rápida de Fourier (FFT)

Funções não-periódicas podem ser representadas pela transformada rápida de Fourier que decompõe um sinal em suas componentes elementares seno e cosseno. A FFT é intensivamente utilizada no estudo de problemas que envolvem vibrações aleatórias e seus respectivos parâmetros (INMAN, 2008).

Ainda segundo Inman (2008), a FFT é uma consideração da série de Fourier de uma função não periódica. A FFT de uma função $X(t)$ é representada por $X(\omega)$ e é definida pela equação 3.1.

$$X(\omega) = \int_{-\infty}^{\infty} X(t)e^{-j\omega t} dt \quad (3.1)$$

A equação 3.1 transforma o sinal $X(t)$ de uma função temporal para uma função de frequência ω , onde $j = \sqrt{-1}$.

A análise de um sinal no domínio do tempo é, muitas vezes, complexa e imprecisa. Dessa forma, a FFT possibilita a visão do sinal por uma perspectiva diferente, facilitando e acrescentando confiabilidade à análise deste sinal.

3.3.2 Transformada de Fourier de Tempo Curto (STFT)

A transformada rápida de Fourier é uma ferramenta poderosa para analisar um sinal que mantém suas características constantes ao longo do tempo, ou seja, a aplicação dessa técnica exige que o sinal seja estacionário e seu espectro de frequência não varie com o tempo. Dessa forma, pode-se concluir que a FFT apresenta resolução apenas no domínio da frequência.

Assim, em uma análise de sinal onde as informações temporais são importantes, a FFT não produz resultados eficientes. Dessa forma, com a necessidade de estudar sinais transientes (não

estacionários), uma adaptação da FFT foi desenvolvida através da aplicação de uma janela deslizante no tempo, de modo a apresentar resolução no tempo e na frequência. Essa técnica de análise por aplicação de janelas não variáveis é chamada de Transformada Rápida de Fourier de Tempo Curto (STFT) (LYRA, 2019).

Nesse sentido, a STFT se caracteriza por ser uma ferramenta matemática que possibilita a extração de informações temporais de um sinal em uma determinada frequência. Portanto, por meio da aplicação da STFT, é possível identificar a variação temporal de características de uma determinada frequência do sinal (SEO et. al., 2018).

O princípio de funcionamento da STFT é a utilização de uma janela de tempo, pequena o bastante para que o sinal seja considerado estacionário nesse espaço de tempo, que desliza em todas as frequências do sinal (CHEN et. al., 2019). A STFT é obtida através da equação 3.2.

$$X(m, f) = \sum_{t=-\infty}^{\infty} x(t)w_2(t-m)e^{-2\pi f t j} \quad (3.2)$$

Onde, $x(t)$ é o sinal no domínio do tempo, $w_2(t-m)$ é a função da janela centrada no tempo m , f representa o intervalo de frequência, e $j = \sqrt{-1}$.

3.3.3 Transformada Wavelet

Transformadas wavelet consistem em decompor um sinal através de uma função base (Wavelet mãe) com seus devidos fatores de escalonamento e deslocamento (LYRA, 2019). Através dessa técnica, é possível realizar uma análise no domínio tempo-frequência em multirresoluções de um sinal não-estacionário.

Uma comparação bastante comum é estabelecida entre a wavelet e a STFT, pois ambas utilizam técnicas de janelamento para a análise do sinal. A grande diferença reside no conceito de escala do janelamento do sinal, a STFT presume a estacionariedade do sinal em um determinado intervalo de tempo e assume um valor fixo de escala para todo o sinal. Dessa forma, a STFT apresentará uma boa resolução no domínio do tempo ou da frequência, mas não apresentará uma boa resolução em ambos os domínios de forma simultânea. Por outro lado, a técnica da Wavelet apresenta um janelamento com variação de escalas, permitindo uma boa resolução nos domínios do tempo e da frequência simultaneamente (ALMEIDA, 2010).

Para a análise das transformadas wavelet, a localização, no domínio do tempo, das componentes de alta frequência de um sinal é realizada através de escalas menores de tempo, já para a análise de baixas frequências, utiliza-se escalas maiores (RIELLA, 2008). Uma demonstração das características de escala da STFT e da Wavelet é representada pela figura 3.6.

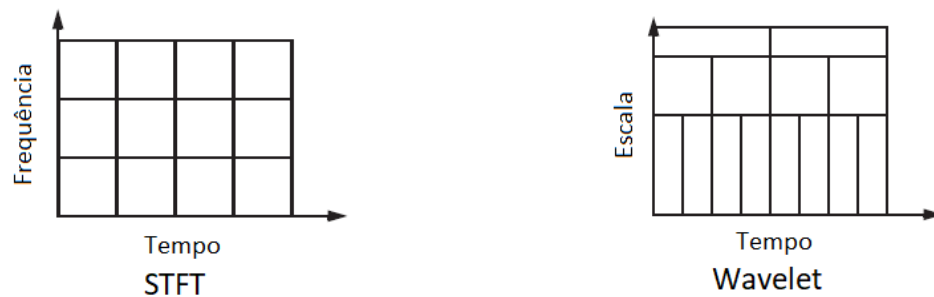


Figura 3. 6 - Análise Comparativa das Características de Escala da STFT e da Wavelet. Fonte: Riella (Adaptado), 2008

Por apresentar um ganho significativo na análise de descontinuidades de sinais em relação à STFT, as transformadas wavelet, em suas diferentes formas, vêm sendo utilizadas em diversas aplicações da ciência moderna, as principais são: análise de defeitos em máquinas; estudos relacionados ao meio ambiente (índices pluviométricos, por exemplo); e pesquisas na área médica (análises de imagem e som, em sua maioria). Basicamente, as transformadas wavelet podem ser estudadas de forma contínua ou discretizada, ambas as formas serão abordadas adiante.

3.3.4 Transformada Wavelet Contínua (CWT)

A Transformada Wavelet Contínua (CWT) de um sinal é dada através de uma operação de convolução entre o sinal no domínio do tempo e a Wavelet mãe (Wavelet original) de uma determinada família (YAN et. al., 2013). A CWT pode ser obtida através da equação (3.3).

$$CWT(s, \tau) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{\infty} x(t) \psi\left(\frac{t-\tau}{s}\right) dt \quad (3.3)$$

Onde ψ representa a função Wavelet mãe, τ é o coeficiente de translação responsável pelo deslocamento da wavelet no eixo do tempo., e s é o parâmetro de escala, ocasionando a dilatação ou contração da função wavelet.

Analisando a equação (3.3) é possível perceber a similaridade entre a transformada wavelet contínua e a transformada de Fourier de tempo curto com a diferença da existência de um fator de translação e, principalmente, a existência de um parâmetro de escala variável (RIELLA, 2008).

Apesar dos ganhos relacionados à análise de sinais, a CWT apresenta uma grande quantidade de dados ao calcular todas as escalas que variam de acordo com o sinal. Portanto, muitas vezes a CWT se torna computacionalmente inviável (PENG; TSE e CHU, 2005). Diante dessa problemática, recorre-se à discretização da transformada wavelet.

3.3.5 Transformada Wavelet Discreta (DWT)

A utilização da CWT gera informações redundantes devido à alteração contínua dos parâmetros de escala e translação. Por vezes, a redundância desses parâmetros é desejada, porém, geralmente, o aumento considerável da carga computacional da transformada contínua não apresenta vantagens significativas para o estudo do sinal. É nesse contexto que a discretização da transformada se faz útil, a DWT é uma forma eficaz de enfatizar a eficiência computacional sem apresentar perdas significativas para a análise do sinal (YAN et. al, 2013).

A DWT permite definir variações espaçadas dos parâmetros de translação e escala sem comprometer a capacidade de análise do sinal. A determinação dos parâmetros de escala, s , e de translação, τ , é dado pelas equações (3.4) e (3.5), respectivamente.

$$s = s_o^j \tag{3.4}$$

$$\tau = k\tau_o s_o^j \tag{3.5}$$

Nessas equações, s_o e τ_o são os passos de escala e translação responsáveis pela variação de s e τ . As variáveis j e k são valores inteiros e positivos e são parâmetros discretos. Com esses parâmetros definidos, é possível chegar a uma equação matemática (3.6) que representa a Transformada Wavelet Discreta.

$$DWT(m, n) = \frac{1}{\sqrt{|s_o^j|}} \int_{-\infty}^{\infty} x(t) \psi\left(\frac{t - k\tau_o s_o^j}{s_o^j}\right) dt \quad (3.6)$$

Segundo Almeida (2010), uma escolha criteriosa dos valores de s_o e τ_o é necessária para que a DWT tenha um bom desempenho. Por questões de simplicidade, é comum utilizar valores de $s_o = 2$ e $\tau_o = 1$. Diz-se, nesse caso, que a wavelet é diádica (potência na base 2) e ortonormal (formação de bases ortonormais) pelas famílias de Wavelet expandidas.

Ao aplicar a DWT, o sinal pode ser decomposto em várias componentes de alta frequência (detalhes) e de baixa frequência (aproximações). Dessa forma, utilizando filtros passa-baixa e passa-alta, é possível obter informações de baixa e alta frequências do sinal, respectivamente (KONAR e CHATTOPADHYAY, 2014). Este procedimento de decomposição do sinal em diferentes faixas de frequências recebe o nome de análise wavelet em multirresolução (AWM) e foi proposto por Mallat e Meyer em 1986. A AWM pode ser representada pela figura 3.7.

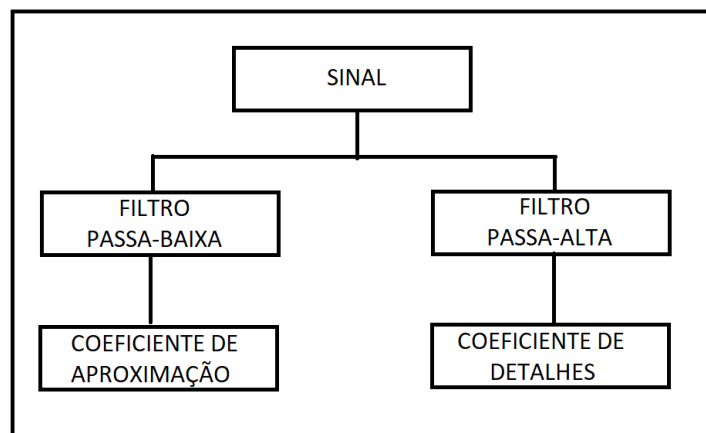


Figura 3. 7 - Representação da Análise Wavelet em Multirresolução. Fonte: Almeida (Adaptado), 2010.

O processo de análise wavelet em multirresolução consiste em decompor o sinal diversas vezes em variáveis de alta frequência (detalhes) e de baixa frequência (aproximação). Nesse processo, a cada decomposição realizada, os coeficientes apresentam a metade da dimensão do sinal que os originou.

Na AWM, os coeficientes A_1 e D_1 são as componentes de aproximação e detalhe do primeiro nível, respectivamente. Para o segundo nível, as componentes são representadas por A_2 e D_2 e assim sucessivamente. A figura 3.8 mostra a decomposição de um sinal em três níveis; porém, esse procedimento pode ser realizado diversas vezes, dependendo apenas da necessidade da análise.

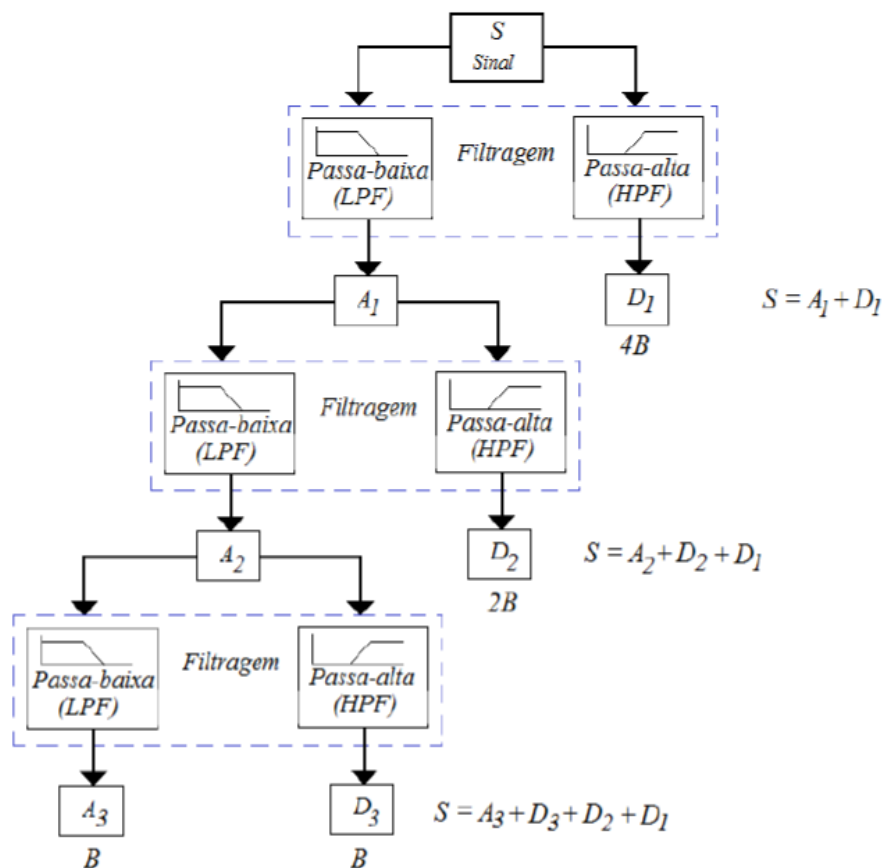


Figura 3. 8 - Decomposição de um Sinal em Três Níveis. Fonte: Lyra (Adaptado), 2019.

Ao analisar o processo de decomposição, mostrado na figura 3.8, verifica-se que o sinal original pode ser reconstruído por meio da soma direta dos detalhes com as aproximações, sendo essa uma das principais características da análise em multirresolução.

3.3.5.1 As Famílias Wavelet

A escolha de uma Wavelet-mãe é determinada através da natureza do sinal e de sua aplicação e, dependendo de suas características, ela otimiza e facilita a análise do sinal (LIMA, 2019).

A primeira família de Wavelet – Wavelet Haar – foi desenvolvida por Alfréd Haar em 1909 e é considerada a mais simples das famílias conhecidas (CAMARGO, 2011). A função da Wavelet de Haar pode ser representada por um grupo de ondas quadradas e ortogonais com magnitude de ± 1 em um determinado intervalo e zero nos demais pontos da função (MALLAT, 2009). A figura 3.9 fornece uma representação gráfica da Wavelet de Haar.

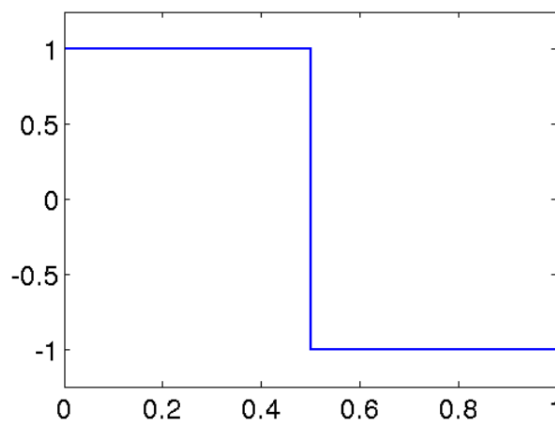


Figura 3. 9 - Wavelet de Haar. Fonte: Almeida (Adaptado), 2010.

Uma outra família tradicional de wavelet bastante utilizada é a família Daubechies, que recebeu esse nome por ter sido desenvolvida por Ingrid Daubechies na década de 80. Em geral, essa família apresenta uma maior efetividade que as wavelets de Haar por apresentar as seguintes características:

- 1- **Suportes compactos:** O domínio da função é diferente de zero em uma extensão finita e igual a zero em todo o resto. Essa propriedade permite a análise local de um determinado sinal.
- 2- **Invariante:** Essa característica da família Daubechies permite a identificação de pequenas mudanças no sinal (HEMANTH et. al., 2018).

- 3- **Momentos nulos:** Quanto maior o número de momentos nulos de uma wavelet, mais suave ela é, ou seja, uma maior quantidade de momentos nulos de uma Wavelet aumenta a probabilidade de reconstrução perfeita do sinal decomposto (UZINSKI, 2013). A família Daubechies de wavelet possui o maior número de momentos nulos para um dado suporte, o que aumenta a eficiência dos filtros de decomposição e melhora a qualidade do suporte (RIELLA, 2008).

A figura 3.10 exemplifica uma Wavelet-mãe da família Daubechies.

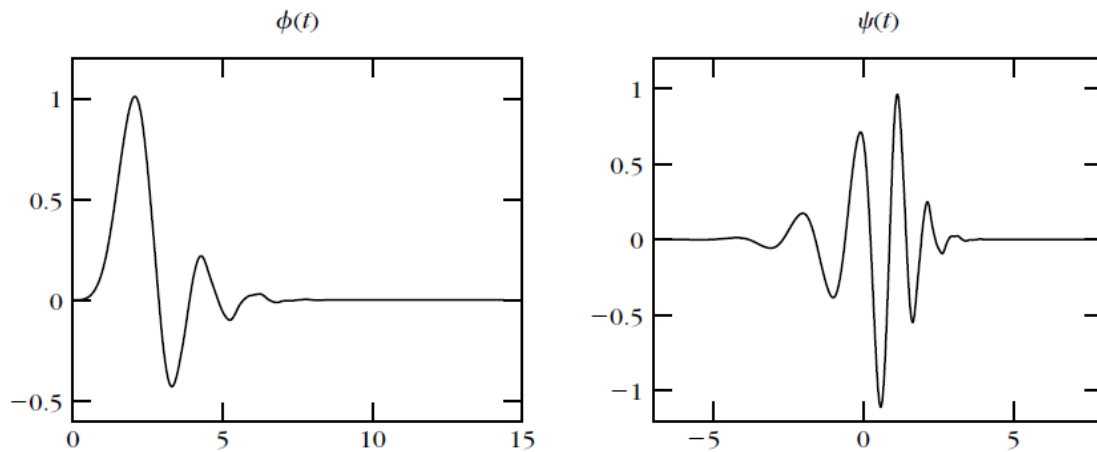


Figura 3. 10 - Representação das funções escala (esquerda) e wavelet-mãe (direita) da família Daubechies. Fonte: Mallat, 2009.

Para aplicações numéricas, o professor R. Coifman, em parceria com I. Daubechies, desenvolveu as coifelets (MALLAT, 2009). Portanto, pode-se dizer que essa família é uma adaptação das wavelets Daubechies e, assim, são compactas e ortogonais. Além disso, quando comparadas com as wavelets Daubechies, as coifelets são consideradas simétricas. A figura 3.11 mostra as funções escala e wavelet da família Coiflet.

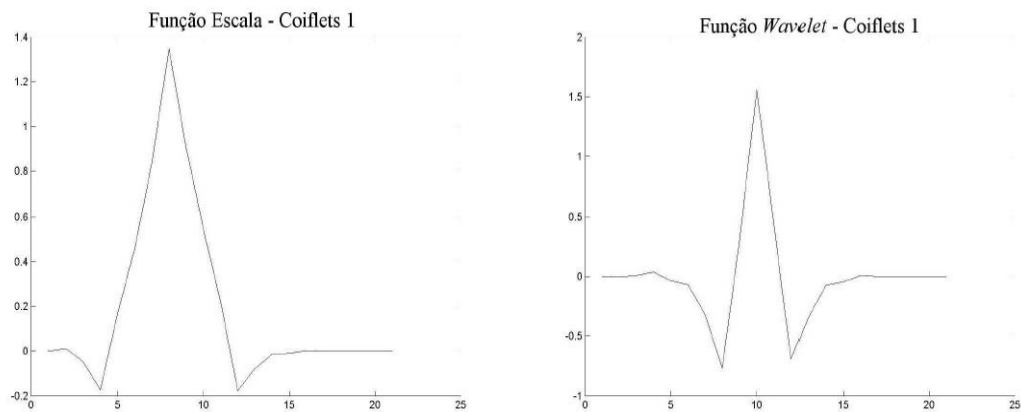


Figura 3. 11 - Funções escala e wavelet da família Coiflet de ordem 1. Fonte: Riella (Adaptado), 2008.

Com o objetivo de resolver o problema da assimetria da wavelet Daubechies, a symmlet foi também desenvolvida por I. Daubechies a partir de filtros simétricos visando obter uma máxima simetria às bases de funções ortonormais. Portanto, a symmlet apresenta características parecidas com as encontradas nas wavelets Daubechies (RIELLA, 2008). A figura 3.12 mostra as funções escala e wavelet da família Symmlet.

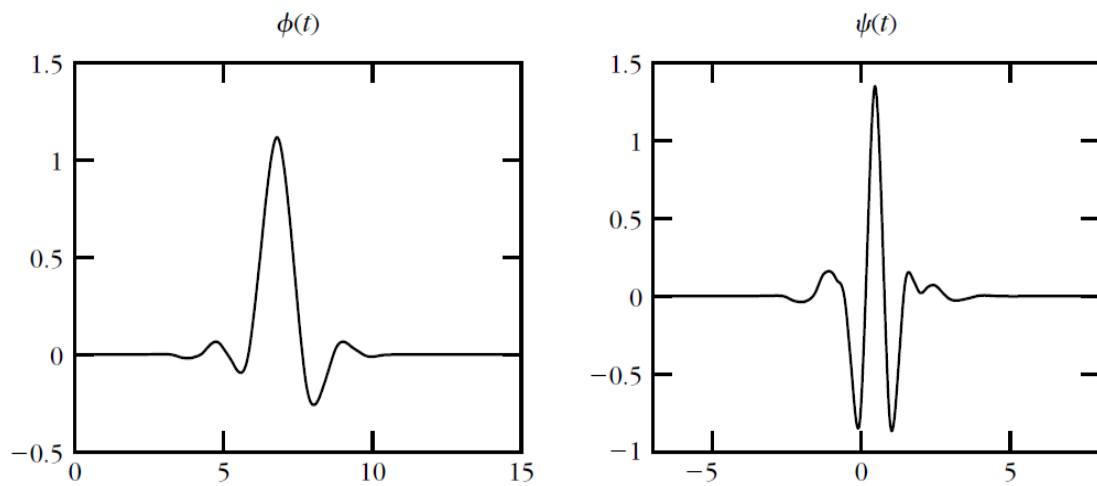


Figura 3. 12 - Funções Escala e Wavelet da família Symmlet. Fonte: Mallat (Adptado), 2009.

3.3.6 Análise Mel-Cepstral

A análise mel-cepstral de sinais tem sido aplicada por diversos pesquisadores com diferentes objetivos. Nesse sentido, Ittichaichareon et. al. (2012) conseguiram reconhecer a voz de indivíduos por meio da aplicação dos coeficientes Mel-Cepstrais (MFCC). Além dessa área de aplicação, os MFCCs têm sido utilizada para identificar patologias laríngeas, como mostra o trabalho de Ribeiro et. al. (2014).

A análise mel-cepstral é extensivamente usada para a extração de características no reconhecimento da voz por sua capacidade de identificar como o aparelho auditivo humano interpreta as diferentes frequências. Para isso, os MFCCs utilizam a escala Mel para analisar as diferentes frequências do espectro.

Em geral, de acordo com Winursito et. al. (2018), a obtenção dos coeficientes cepstrais são calculados em 6 etapas, são elas:

- 1- Pré-Ênfase: Atenua os componentes de altas frequências, por meio do seguinte filtro de passa-alta:

$$y(n) = s(n) - \alpha s(n - 1) \quad (3.7)$$

Onde, $y(n)$ é sinal compensado, $s(n)$ é o sinal de entrada e α é uma constante entre 0,9 e 1. Quanto mais próximo de 1, maior será a ênfase atribuída por α nas altas frequências.

- 2- Janelamento: O sinal é dividido em pequenas partes de comprimento aproximado entre 20 e 40 ms. O tamanho das janelas é o bastante para considerar o sinal estacionário nesse espaço de tempo e, a partir disso, pode-se aplicar a transformada de Fourier. Além disso, tal janelamento condiz com a duração dos fonemas e, portanto, possibilita o estudo da voz. A janela de Hamming é a mais aplicada nesse processo (CERÓN e BADILLO, 2011).

- 3- Transformada Discreta de Fourier (DFT): A extração espectral de informações é dada por meio da aplicação da transformada discreta de Fourier. Nesse sentido, a DFT fornece a energia contida no sinal em suas diferentes frequências.
- 4- Banco de Filtros Mel: Aplica-se os filtros triangulares em escala mel para o mapeamento da forma como o sistema auditivo humano interpreta o som. Nesse banco, os filtros triangulares passa-banda são sobrepostos e as frequências centrais são equidistantes. Além disso, essa escala é linear até 1000 Hz e logarítmica para frequências superiores (RIBEIRO et. al., 2014). Por fim, A equação (3.8) apresenta a função de mapeamento da frequência para uma escala de frequências Mel, onde F_{linear} representa a frequência linear em Hertz.

$$F_{mel} = 2595 \log_{10} \left(1 + \frac{F_{linear}}{700} \right) \quad (3.8)$$

- 5- Logaritmo: Calcula-se o logaritmo da energia de saída de cada filtro de modo a diminuir ou retirar a sensibilidade das variações de potência devido às diferenças de distâncias entre quem está falando e o microfone. Esse processo resulta na formação dos coeficientes logarítmicos de energia (CERÓN e BADILLO, 2011).
- 6- Transformada Discreta do Cosseno: Nessa etapa, os coeficientes obtidos no estágio anterior são ortogonalizados por meio da transformada discreta do cosseno (DCT). Esse procedimento tem o objetivo de comprimir os dados espectrais e formar o vetor cepstral. A equação (3.9) mostra a formulação matemática da DCT. Nessa expressão, C_i é o i -ésimo MFCC, M representa o número de coeficientes cepstrais e X_k é o coeficiente logarítmico de energia do k -ésimo filtro.

$$C_i = \sum_{k=1}^M X_k \cos \left[\left(k - \frac{1}{2} \right) \frac{\pi}{M} \right], \quad i = 1, 2, 3 \dots, M \quad (3.9)$$

Por fim, a figura 3.13 apresenta, de forma resumida, as etapas de obtenção dos coeficientes cepstrais.

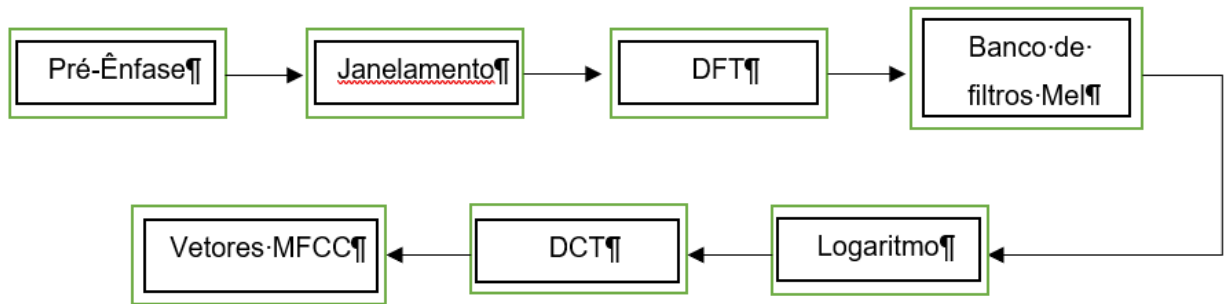


Figura 3. 13 - Etapas de obtenção dos MFCCs.

3.4 REDES NEURAIS

Redes Neurais Artificiais (RNAs) são modelos matemáticos baseados em redes neurais biológicas, onde unidades básicas de processamento, chamadas de neurônio, trabalham em paralelo e aprendem com a experiência de suas iterações (SHALEV-SHWARTZ; BEN-DAVID, 2014).

Segundo Patterson e Gibson (2017), o processo de aprendizagem das redes neurais consiste em um sistema de pesos das informações compartilhadas entre os neurônios. Na medida que esses pesos são atribuídos às novas informações que chegam ao neurônio, os parâmetros da rede neural são estabelecidos.

Ainda segundo o mesmo autor, o comportamento de uma rede neural é determinado por sua arquitetura, que por sua vez é definida pelos seguintes parâmetros:

- 1- Número de neurônios;
- 2- Número de Camadas;
- 3- Tipo de conexões entre as camadas.

Assim, esta seção destaca todo o processo de funcionamento de uma Rede Neural Artificial, desde sua unidade básica de processamento até sua arquitetura final.

3.4.1 Modelo de um Neurônio Artificial

O neurônio artificial, estrutura básica de uma RNA, foi proposto por Mcculloch e Pitts em 1943 e teve como base o funcionamento do neurônio biológico, estrutura básica do sistema nervoso humano.

A principal função do neurônio biológico é enviar e receber impulsos, podendo eles serem excitatórios ou inibitórios. Somam-se aos impulsos valores de pesos diferentes e, após isso, esses valores finais são enviados aos neurônios vizinhos. A liberação de estímulos e produtos bioquímicos por parte do organismo vai depender da força do impulso gerado, ou seja, depende da soma do impulso gerado com o peso associado a ele (WILAMOWSK; IRWIN, 2011).

Assim como acontece com os neurônios biológicos, as conexões entre os neurônios artificiais estão associadas a pesos. Nesse sentido, os pesos podem ser excitatórios (positivos) ou inibitórios (negativos), sendo ajustados no processo de aprendizagem da RNA (FACELI et al., 2011).

O neurônio artificial é dividido em quatro partes; são elas: sinapses, bias (polarização), somador e função de ativação (RIELLA, 2008). Assim, a estrutura do neurônio proposta por Mcculloch e Pitts é mostrada na figura 3.14.

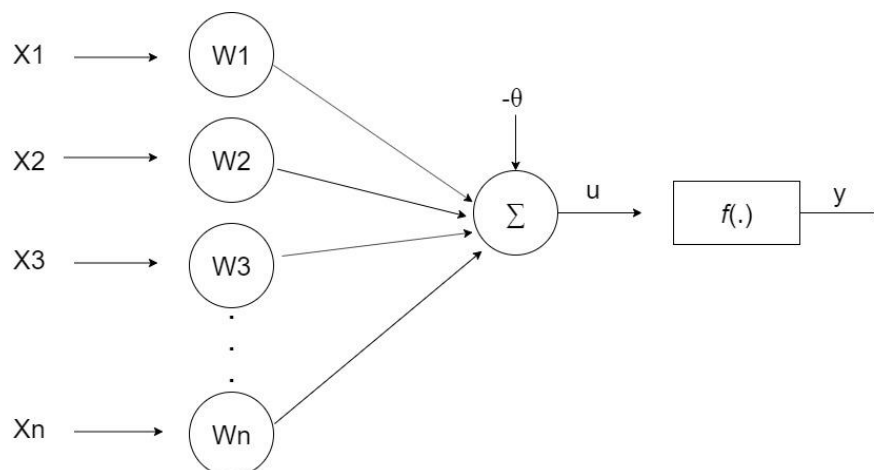


Figura 3. 14 - Neurônio Artificial. Fonte: Patterson e Gibson (Adaptado), 2017.

Onde, $X_1, X_2, X_3, \dots, X_n$ representam os estímulos externos recebidos pelo neurônio (sinais). $W_1, W_2, W_3, \dots, W_n$ são os pesos associados aos sinais de entrada, o que permite quantificar a relevância de cada estímulo. O somador, representado pelo símbolo Σ , realiza uma combinação linear entre os sinais de entrada com os seus pesos. O bias, θ , especifica o patamar de

ativação para que o resultado do somador possa produzir um valor de disparo para outro neurônio na saída. O símbolo u representa o potencial de ativação do neurônio e é resultado da diferença entre a combinação linear, realizada pelo somador, e o bias. Por fim, a função de ativação $f(.)$ gera o sinal de saída (y) do neurônio e tem como objetivo limitar essa saída em um intervalo conhecido (LYRA, 2019). Assim, os parâmetros de entrada e saída do neurônio podem ser descritos matematicamente de acordo com as equações 3.10 e 3.11, respectivamente.

$$u_k = \sum_k (W_k X_k - \theta_k) \quad (3.10)$$

$$y_k = f(u_k) \quad (3.11)$$

3.4.2 Funções de Ativação

As funções de ativação são utilizadas com o intuito de introduzir a não linearidade nos dados de entrada da rede, fazendo com que a RNA seja capaz de realizar tarefas mais complexas (PATTERSON e GIBSON, 2017). Dentre os vários tipos de funções de ativação, as que mais se destacam são: função linear, função limiar, função sigmoide e a função softmax.

3.4.2.1 Função Linear

Ao utilizar a função linear, o neurônio realizará uma transformação linear. Dessa forma, a função linear propaga o sinal original. Esse tipo de função é utilizado na camada de entrada da RNA, podendo ser descrita de acordo com a equação 3.12 e representada por meio da figura 3.15 (PATTERSON e GIBSON, 2017).

$$f(X) = WX \quad (3.12)$$

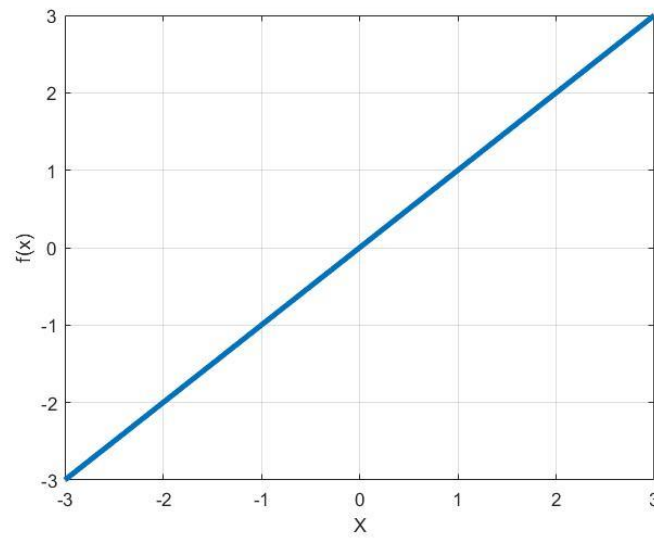


Figura 3. 15 - Representação Gráfica da Função Linear.

3.4.2.2 Função Limiar

A função limiar assume apenas dois valores, 0 e 1. Quando o somatório de todas as entradas, associadas aos respectivos pesos, com o bias não negativo, a saída reproduzida por essa unidade básica assumirá o valor 1, caso contrário, seu valor será 0 (HAYKIN, 2009).

Neste modelo, a retropropagação do erro é impossibilitada, devido à derivada da função limiar ser nula ou inexistente (KRIESEL, 2005).

A função limiar é descrita matematicamente pela equação 3.13.

$$f(u) = \begin{cases} 1 & \text{se } u_k \geq 0 \\ 0 & \text{se } u_k < 0 \end{cases} \quad (3.13)$$

Onde u_k é o valor do somatório das entradas com o bias e pode ser determinado através da equação 3.10. A representação gráfica da função limiar é mostrada na figura 3.15.

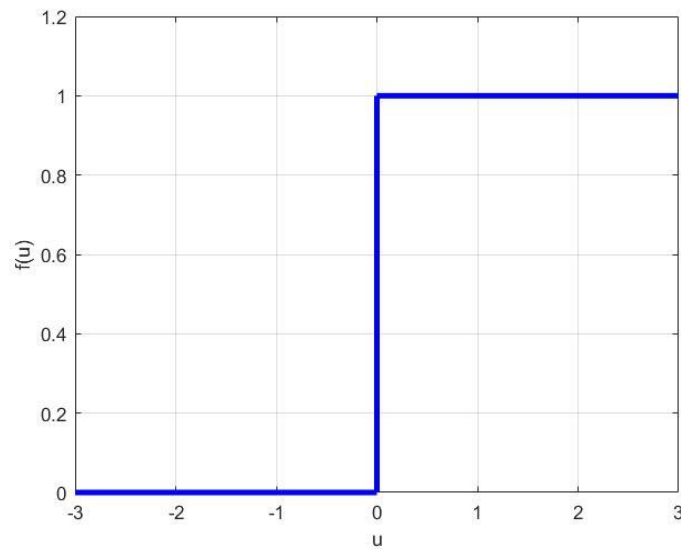


Figura 3.15 – Representação Gráfica da Função Limiar.

3.4.2.3 Função Sigmóide

Ao contrário da função limiar, a função sigmoide é diferenciável em qualquer ponto da curva (KRIESEL, 2005). Segundo Riella (2008), essa função de ativação é a mais utilizada na construção de redes neurais artificiais.

As duas formas mais comuns de utilização da função sigmóide são: função tangencial hiperbólica e função logística.

1- Função Logística:

A forma geral dessa função sigmoidal é definida pela equação 3.14.

$$f(u) = \frac{1}{1+e^{-au}} \quad (3.14)$$

Nesta equação, a é o parâmetro de inclinação da curva, u é o campo local induzido do neurônio. Quanto maior o valor de a , menor será a inclinação da curva da função logística, mostrada na figura 3.16. Sendo assim, no infinito, a função sigmoide se transforma na função limiar.

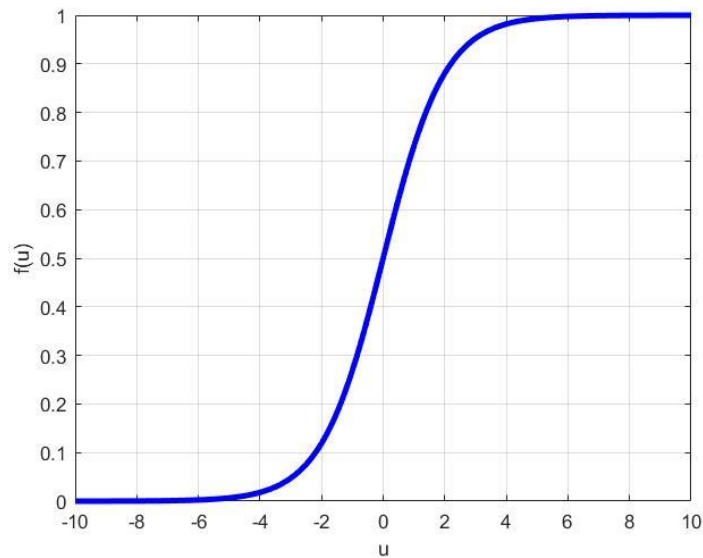


Figura 3. 16 - Representação Gráfica da Função Logística.

2- Função Tangencial Hiperbólica:

A tangente hiperbólica (TANH) é uma outra forma da função sigmóide e representa a razão entre o seno hiperbólico e o cosseno hiperbólico. A TANH é limitada ao intervalo de -1 a 1, enquanto a função logística está limitada ao intervalo 0 a 1. Dessa forma, a grande vantagem da tangente hiperbólica sobre a função logística é que ela pode lidar com valores negativos com facilidade (PATTERSON e GIBSON, 2017).

A forma geral da função tangencial hiperbólica é mostrada através da equação 3.15

$$f(u) = \tanh(u) = \frac{\sinh(u)}{\cosh(u)} \quad (3.15)$$

A função tangencial hiperbólica pode ser representada pela figura 3.17.

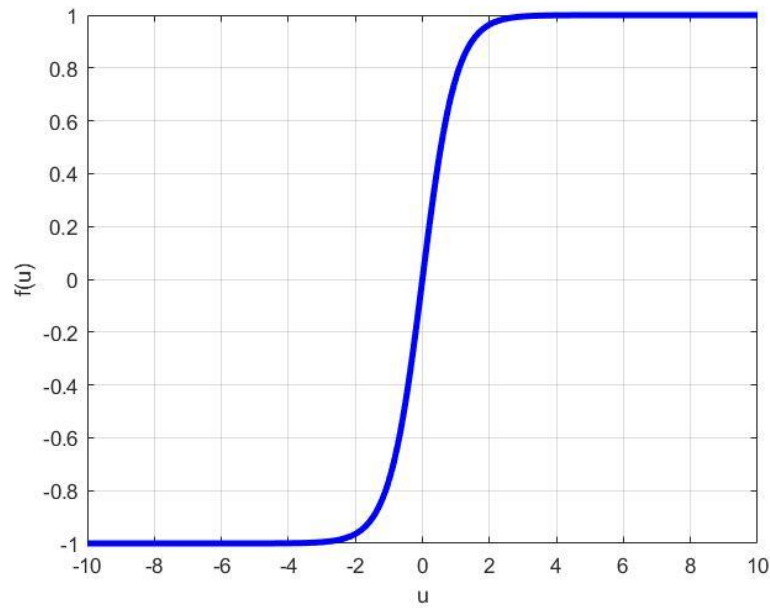


Figura 3. 17 - Representação Gráfica da Função Tangencial Hiperbólica.

3.4.2.4 Função Softmax

A função softmax é frequentemente utilizada na camada de saída de RNAs que objetivam classificar um sinal em múltiplas classes através de uma distribuição de probabilidade (GAO; PAVEL, 2017). Sendo assim, a classe do sinal será aquela em que apresenta um maior valor de probabilidade e o somatório dos valores de todas as classes deverá ser 1. A função softmax é definida pela equação 3.16.

$$\hat{y}_i = \frac{e^{u_i}}{\sum_j e^{u_j}} \quad (3.16)$$

Onde, i é o número de classes e $\hat{y}$ é a distribuição da probabilidade e, portanto, $\hat{y}_1 + \hat{y}_2 + \dots + \hat{y}_i = 1$.

3.4.3. Redes de Camada Simples

As redes de camada simples, ou *perceptrons*, são caracterizadas por apresentar suas entradas diretamente conectadas com as saídas. Essas redes possuem apenas uma matriz de pesos e sua função de ativação é a limiar. A representação gráfica das redes de camada simples é apresentada na figura 3.18 (RIELLA, 2008).

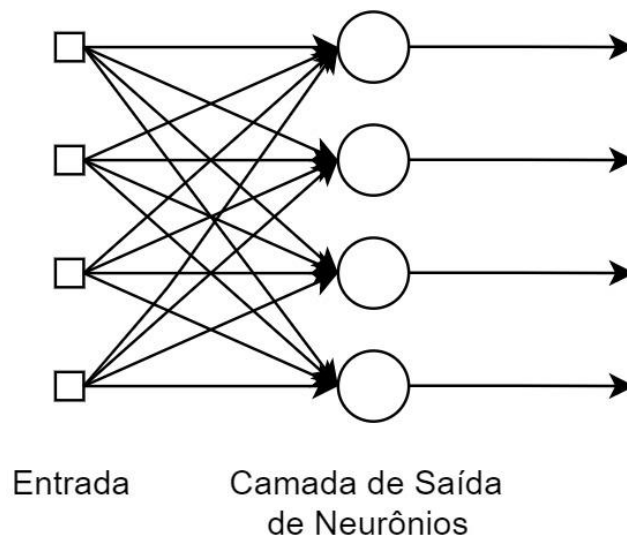


Figura 3. 18 - Rede Neural Artificial de Camada Simples.

Segundo Haykin (2008), a rede perceptron é a rede neural mais simples para a classificação de padrões que se encontram em lados opostos de um hiperplano (padrões linearmente separáveis).

3.4.5 Redes de Múltiplas Camadas

Uma rede neural com duas ou mais camadas de neurônios com pesos ajustáveis é chamada de rede de múltiplas camadas ou perceptron de múltiplas camadas (MLP). A principal vantagem da MLP em relação às redes de camadas simples é sua maior capacidade de interações entre os neurônios.

Dessa forma, toda MLP possui, no mínimo, três camadas, são elas: camada de entrada, camada(s) oculta(s) e camada saída (RIELLA, 2008). A representação gráfica da MLP é mostrada na figura 3.19.

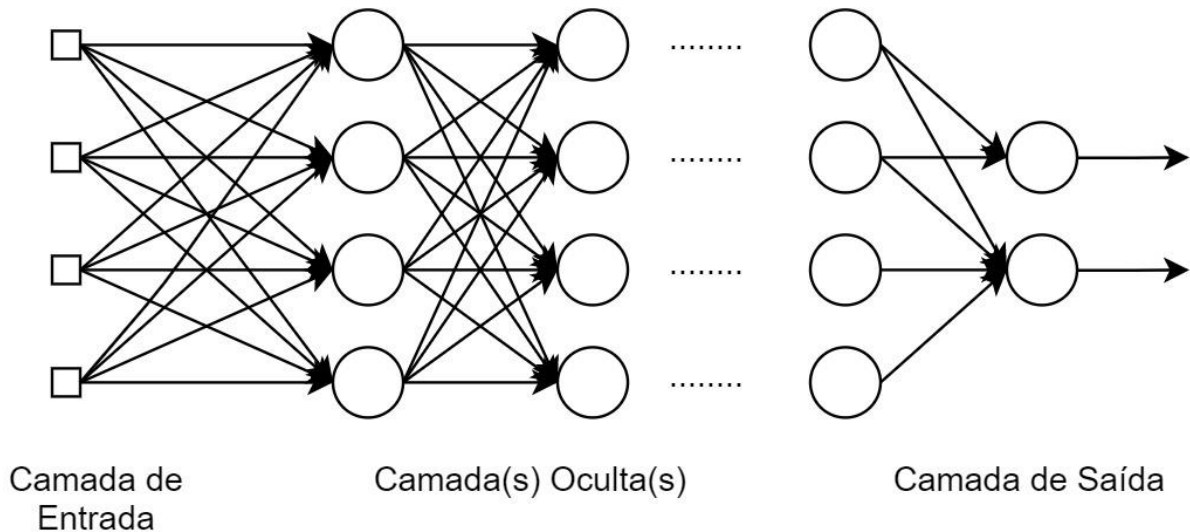


Figura 3. 19 - Rede Neural Artificial de Múltiplas Camadas.

A camada de entrada recebe o estímulo (sinal de entrada) do meio externo que é passado adiante para a(s) camada(s) oculta(s), associado a pesos ajustáveis de acordo com o processo de aprendizagem da rede. A função de ativação da camada de entrada é a linear, isto é, a informação não vai ser modificada e será apenas transferida para o interior da rede. Nas camadas ocultas, a não linearidade é introduzida ao sinal de entrada por meio da função de ativação (HAYKIN, 2009). Por último, a camada de saída da MLP fornece a resposta da rede a partir dos parâmetros obtidos em seu treinamento.

3.4.6 Processo de Aprendizado Da Rede

Em um processo de aprendizagem, o sistema se adapta as mudanças do meio ao qual está inserido. De acordo com Kriesel (2005), o processo de aprendizagem consiste em ajustar os pesos associados ao sinal de entrada de forma a diminuir o erro da resposta da RNA. Assim, o treinamento ocorre através das interações entre neurônios de diferentes camadas. Dessa forma, diz-se que a rede está treinada quando esses pesos são ajustados de forma que o erro encontrado

na saída da rede seja mínimo. Para isso acontecer, o erro deve ser retropropagado até que o alvo seja atingido.

O algoritmo de retropropagação é uma abordagem clássica para o treinamento de redes neurais artificiais. Esse algoritmo é baseado no gradiente descendente que tem como principal objetivo a minimização do erro quadrático apresentado na saída da RNA (DAUMÉ, 2012).

Outro fator importante no processo de treinamento da rede é a taxa de aprendizado que se refere à taxa de ajustes dos parâmetros durante o treinamento. Durante a retropropagação do erro, o gradiente de erro é multiplicado pela taxa de aprendizado para que os pesos de conexão sejam atualizados (PATTERSON e GIBSON, 2017).

3.5 CONCLUSÃO

Nesse capítulo, foi apresentada a teoria necessária para o desenvolvimento desta dissertação.

Começando pela anatomia e funcionamento do coração, foi abordada as principais estruturas e funções desse órgão vital assim como a patologia que é objeto de estudo deste trabalho e seu diagnóstico.

Para o estudo dos sinais, foram apresentadas diversas técnicas de análise de sinais. Por facilitar a extração de informações relevantes, foram abordadas as técnicas de análise nos domínios da frequência e tempo-frequência.

No domínio da frequência, a FFT é a técnica mais utilizada para o estudo dos sinais estacionários. Porém, como os sinais provenientes das vozes humanas são transientes, a aplicação da FFT não é a mais indicada para esse estudo. Com o objetivo de superar essa dificuldade, optou-se por usar as análises no domínio tempo-frequência (STFT, DWT e MFCC) pois essas são capazes de extrair informações de sinais variantes no tempo.

Após a análise dos sinais, as informações extraídas serviram para o desenvolvimento das redes neurais. As RNAs são modelos matemáticos capazes de identificar padrões. Dessa forma, tais modelos são utilizados nesta pesquisa para identificar a ocorrência da IC.

Portanto, o presente capítulo apresenta toda a teoria necessária para o desenvolvimento do aplicativo para auxiliar os profissionais da saúde para o diagnóstico da insuficiência cardíaca.

CAPÍTULO IV

MATERIAIS E MÉTODOS

O capítulo 6 desta dissertação direciona-se à explanação da metodologia utilizada para o desenvolvimento do sistema de auxílio na identificação da IC. Aqui, será abordado o processo de aquisição, seleção, pré-processamento e processamento dos sinais sonoros aqisitados assim como o equipamento utilizado para a gravação das vozes.

4.1 FLUXOGRAMA DA PESQUISA

A figura 4.1 apresenta as etapas para o desenvolvimento da pesquisa.

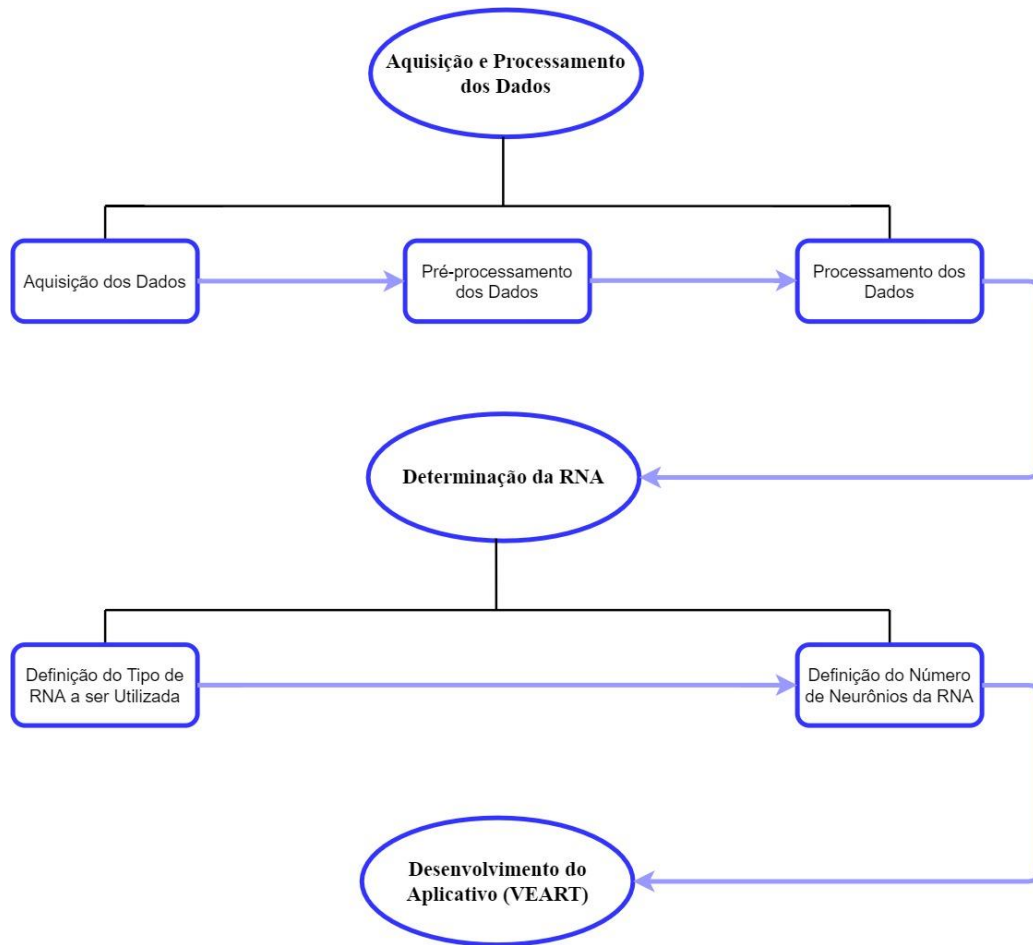


Figura 4. 1 - Fluxograma da Pesquisa.

4.2 AQUISIÇÃO DE DADOS

O processo de aquisição de dados foi dividido em duas etapas:

- 1- Aquisição das vozes de pessoas com IC: Essa fase foi realizada no INCOR (USP) e no hospital metropolitano de João Pessoa, na qual todas as aquisições foram de pacientes com insuficiência cardíaca previamente diagnosticada no ambiente hospitalar.
- 2- Aquisição das vozes de pessoas saudáveis: Nessa etapa, as aquisições foram realizadas com indivíduos saudáveis em ambiente extra hospitalar. Para uma pessoa apresentar o status de saudável, esta não poderia apresentar nenhum sintoma que pudesse ser relacionado com a IC.

Ao realizar a gravação das vozes no INCOR, os dados clínicos (NYHA e FEVE) dos pacientes foram previamente analisados por meio do acesso ao banco de dados do próprio instituto. Além da informação obtida no banco de dados, todo o procedimento experimental foi acompanhado por um profissional de saúde do INCOR. Já no hospital metropolitano, os dados foram aquiridos de pacientes que estavam internados na instituição e que foram previamente diagnosticados com a doença. Além disso, todas as gravações foram realizadas por médicos especialistas em cardiologia.

Durante a etapa de gravação das vozes dos indivíduos saudáveis e não saudáveis, o ambiente foi controlado com o objetivo de diminuir a influência de ruído externo no presente estudo. Além disso, o mesmo equipamento foi empregado na aquisição de todos os dados, o que possibilitou a padronização dos sinais aquiridos.

Ao gravar as vozes, solicitou-se que os participantes da pesquisa repetissem três vezes a seguinte frase:

“É importante cuidar do coração”.

A frase acima foi utilizada por apresentar um baixo grau de dificuldade para memorização, visto que é composta por apenas 5 palavras e sua repetição levou, em média, 7 segundos por gravação. Para a aquisição dos sinais sonoros das vozes, utilizou-se o gravador de voz PX440 da Sony, que tem suas especificações técnicas apresentadas na tabela 4.1 e é apresentado na figura 4.2.

Especificações técnicas	Descrição
Memória Interna	4 GB
Microfone Interno	Estéreo
Formato de Reprodução	mp3
Formato de Gravação	mp3
Taxa de Bits	8/48/128/192 Kbps
Tempo Máximo de Gravação a 128 Kbps	67h 05min
Resposta de Frequência a 128 Kbps	75 – 17000 Hz

Tabela 4. 1 – Especificações técnicas do gravador digital PX440 da Sony.



Figura 4. 2 - Gravador digital PX440 da Sony.

Assim, o gravador PX440 foi utilizado por sua capacidade de gravar as vozes na qualidade requerida pela pesquisa, além de ser de baixo custo e fácil uso.

4.3 CRITÉRIOS PARA PARTICIPAÇÃO NA PESQUISA

Esta pesquisa trata de um estudo piloto em que foram avaliadas as vozes de pacientes portadores de insuficiência cardíaca e comparadas com as vozes de indivíduos saudáveis. Os dados aquistados foram ordenados por sexo e idade dos entrevistados. Para participar da pesquisa, os indivíduos selecionados seguiram os critérios de inclusão e exclusão pré-estabelecidos mostrados na tabela 4.2.

Critérios de exclusão	Critérios de inclusão
Idade inferior a 18 anos	FEVE < 40%
Tabagistas	NYHA II, III, IV
Traumas ou tumores craniocervicais	
Infecção de vias aéreas superiores no momento da gravação	
Alterações tireoidianas descompensadas	

Malformações congênitas craniocervicais	
Cirurgias de laringe ou traqueostomia	
Distúrbio de fala (gagueira, mutismo)	
Gestantes	
Transtornos psiquiátricos	
Doenças Neurodegenerativas	

Tabela 4. 2 - Critérios de Exclusão e Inclusão.

Além disso, os indivíduos que se negaram a assinar o termo de consentimento livre e esclarecido (TCLE) foram excluídos da pesquisa.

4.4 PREPARAÇÃO DOS DADOS

Após a etapa de aquisição, os dados foram organizados em classes de idade que, por sua vez, são separadas por sexo. Essa etapa é importante para que os efeitos da idade e do sexo na voz sejam eliminados na criação da rede neural. Assim, foram criadas duas redes neurais, uma para cada sexo, onde os dados para alimentar as RNAs são provenientes de pessoas, doentes ou saudáveis, com idades situadas nas mesmas classes. As figuras 4.3 e 4.4 mostram o número de pessoas em cada classe de indivíduos com IC e saudáveis, respectivamente, para o sexo masculino utilizado para criar a rede. Por outro lado, as figuras 4.5 e 4.6 mostram as classes para os indivíduos com IC e saudáveis do sexo feminino.

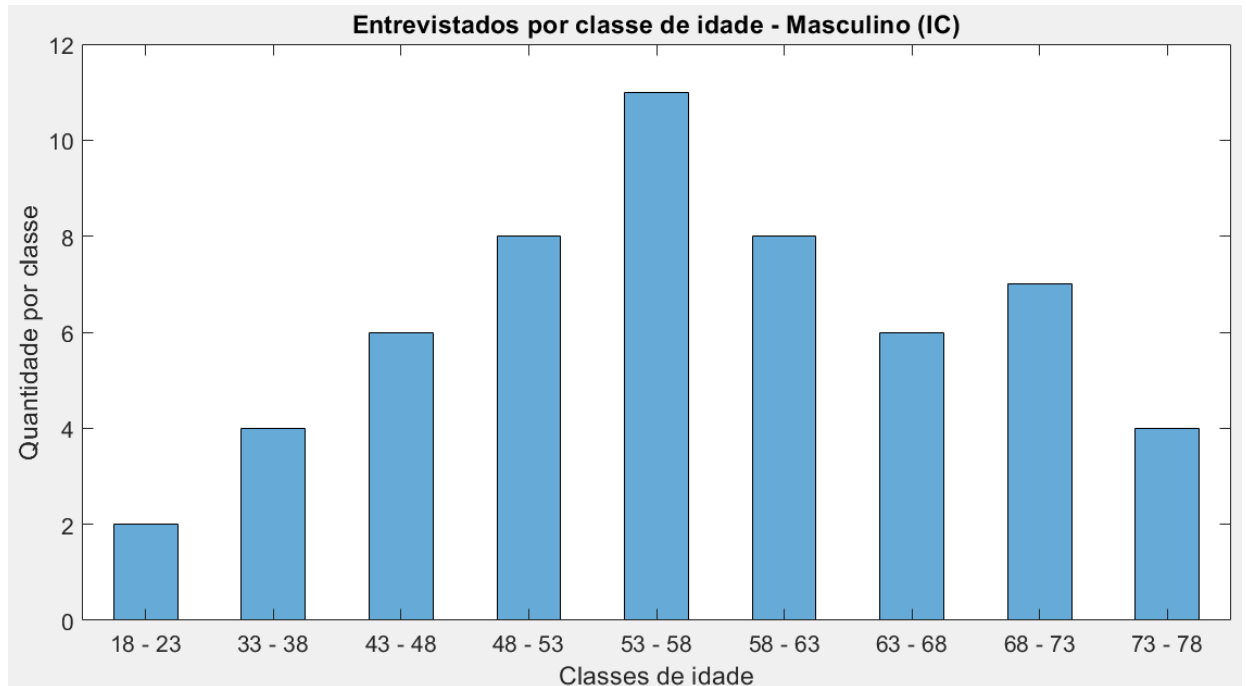


Figura 4. 3 - Distribuição das classes por idade – Masculino (IC).

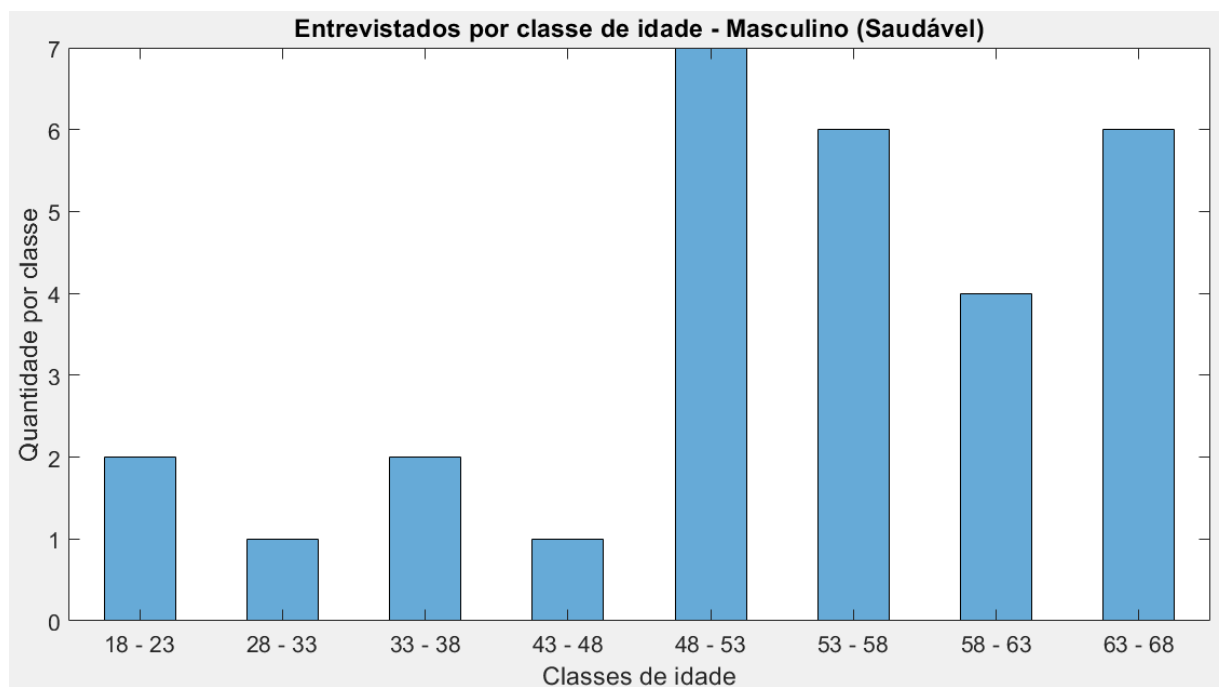


Figura 4. 4 - Distribuição das classes de idade – Masculino (Saudável).

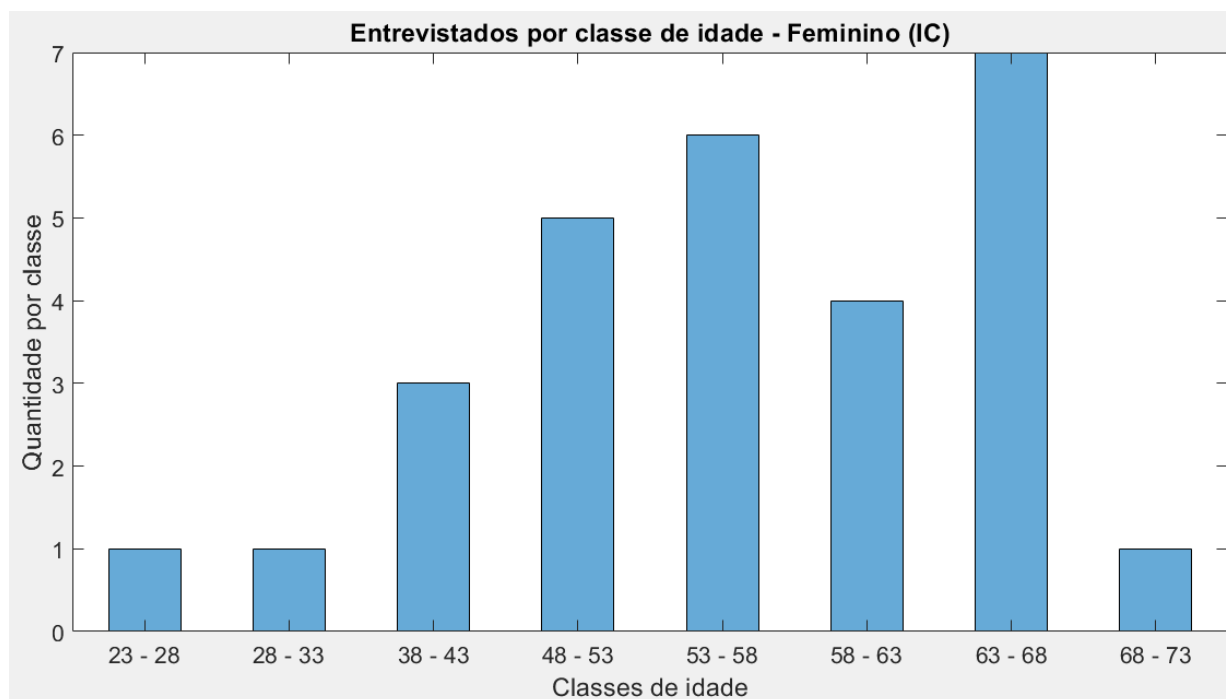


Figura 4. 5 - Distribuição das classes de idade – Feminino (IC).

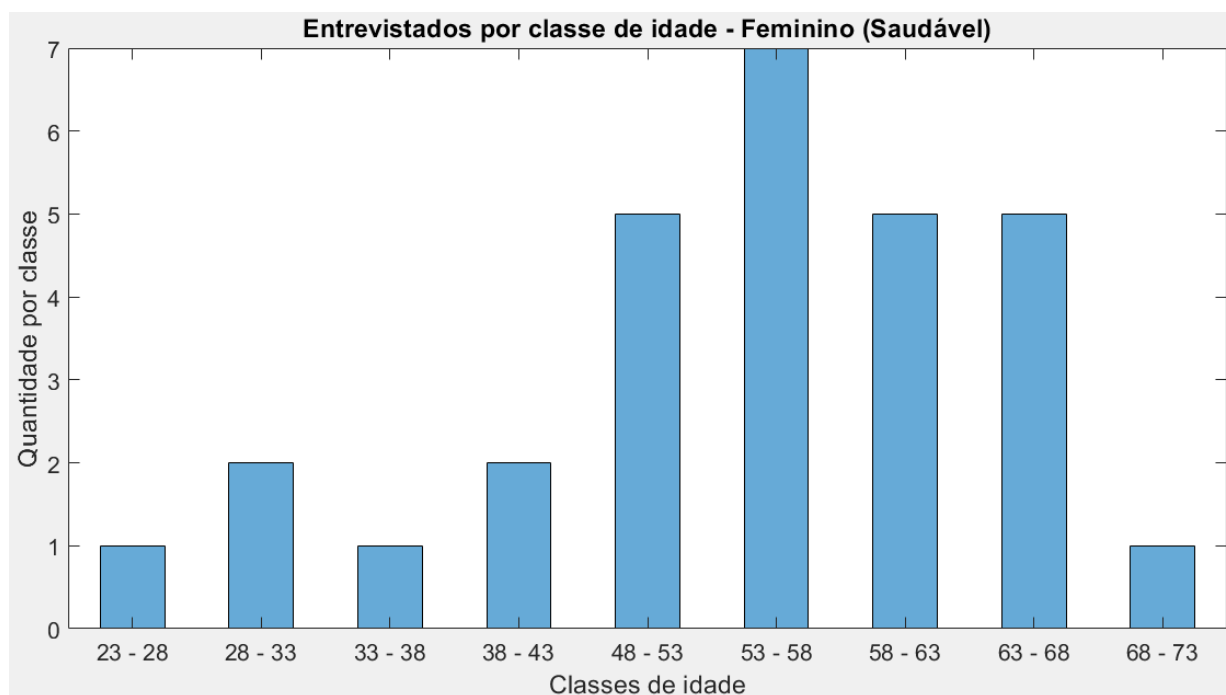


Figura 4. 6 - Distribuição das classes de idade – Feminino (Saudável).

Para os indivíduos com insuficiência cardíaca, as amostras foram feitas de forma aleatória. Assim, as classes de homens e mulheres com IC (figuras 4.3 e 4.5, respectivamente) foram formadas à medida que os pacientes foram entrevistados no INCOR e no hospital metropolitano. Por outro lado, as classes dos indivíduos saudáveis – homens e mulheres – foram formadas com a

intenção de repetir a mesma distribuição das classes dos pacientes com IC diagnosticada. Essa metodologia de organização dos dados teve a intenção de anular os efeitos da idade na voz dos indivíduos ao serem desenvolvidas as RNAs. Por fim, a tabela 4.3 apresenta a distribuição da quantidade de pessoas para cada classe de idade.

Estado de Saúde	Saudáveis		Doentes	
Classes	Homens	Mulheres	Homens	Mulheres
18 - 23	2	0	2	0
23 - 28	0	1	0	1
28 - 33	1	2	0	1
33 - 38	2	1	4	0
38 - 43	0	2	0	3
43 - 48	1	0	6	0
48 - 53	7	5	8	5
53 - 58	6	7	10	6
58 - 63	4	5	8	4
63 - 68	6	5	6	7
68 - 73	0	1	7	1
73 - 78	0	0	4	0

Tabela 4. 3 – Análise de classe.

Assim, por meio da tabela 4.3, percebe-se que a análise de classes de idades possibilita a aplicação da mesma quantidade de indivíduos em cada classe utilizados para o desenvolvimento das RNAs.

4.5 PROCESSAMENTO DOS DADOS

Por padrão do equipamento utilizado na aquisição de dados, os sinais obtidos estão no domínio do tempo e salvos em formato MP3. No entanto, os sinais no domínio do tempo caracterizam-se por sua dificuldade de extração de parâmetros relevantes (frequência fundamental, dados estatísticos de determinadas faixas de frequência, entre outros) para que um estudo minucioso seja realizado sobre eles. Dessa forma, as técnicas de análise de sinais abordadas no capítulo 3 desta dissertação – DWT, análise estatística, FFT e análise mel cepstral – foram empregadas de modo a extrair características do fenômeno que está sendo estudado, tais como amplitude do sinal, média e frequência fundamental, por exemplo; além de possibilitar a análise

em multirresolução dos sinais, no caso da aplicação da DWT. Os dados obtidos nessa etapa da pesquisa serviram como *input* para o desenvolvimento da RNA a ser utilizada na determinação do padrão da IC.

4.5.1 Transformada Wavelet Discreta (DWT)

A análise dos sinais iniciou-se com a aplicação da DWT ao sinal com o objetivo de realizar uma análise em multirresolução dos dados. Assim, cada faixa de frequência da voz gravada poderá ser estudada com maior precisão e, caso essa faixa seja relevante para o estudo, suas características serão extraídas e usadas como entrada para a rede neural.

Ao aplicar a DWT nos sinais, os dados foram decompostos em 11 detalhes, isto é, faixas de frequências. Nesse sentido, a tabela 4.4 correlaciona os detalhes da DWT nas suas respectivas faixas de frequência.

Detalhe	Faixa de Frequência (Hz)
1	11025 – 22050
2	5512,50 – 11025
3	2756,25 – 5512,50
4	1378,13 – 2756,25
5	689,06 – 1378,13
6	344,53 – 689,06
7	172,27 – 344,53
8	86,13 – 172,27
9	43,07 – 86,27
10	21,53 – 43,07
11	10,77 – 51,53

Tabela 4. 4 - Faixa de Frequência dos Detalhes.

Para exemplificar o modelo de decomposição aplicado, a figura 4.7 apresenta um sinal sonoro de uma voz feminina em multirresolução decomposto em 4 detalhes. Analisando a figura 4.7, o sinal original está na primeira linha, seguido pelos sinais dos detalhes 1, 2, 3 e 4, respectivamente.

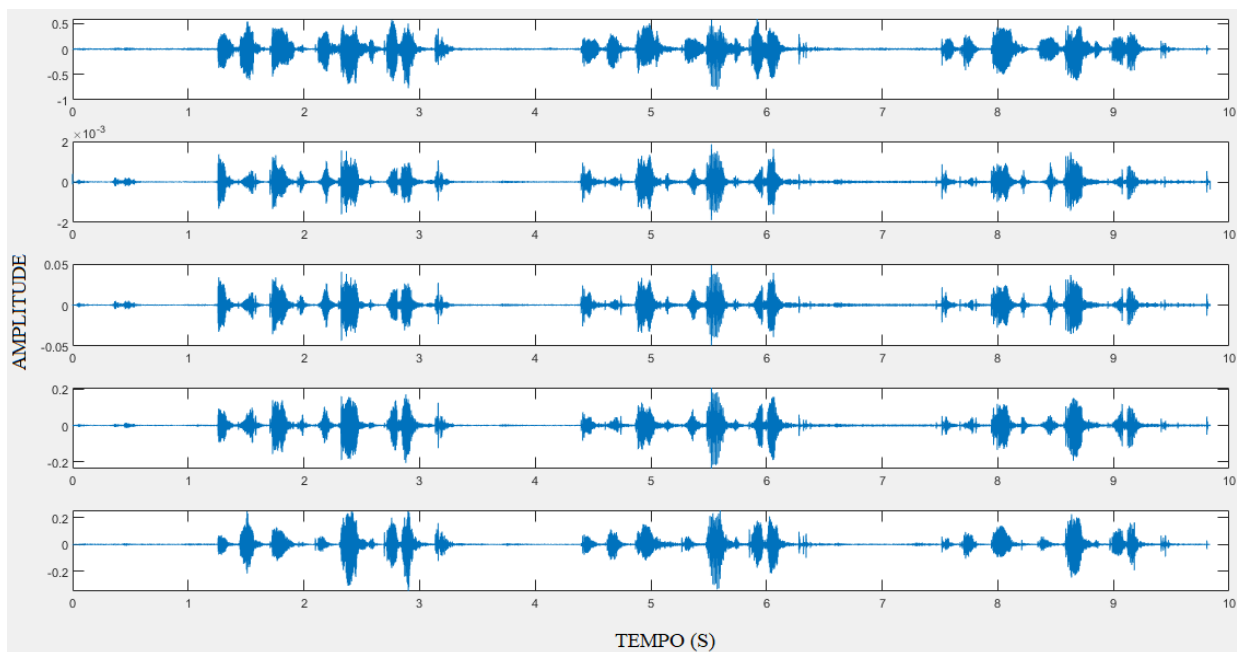


Figura 4. 7 – Análise em multirresolução do sinal sonoro de uma voz feminina.

Apesar da figura 4.7 apresentar uma visão geral da análise em multirresolução de um sinal, a sua visualização fica comprometida pelo excesso de informação nela contida. Assim, as figuras 4.8 a 4.11 possibilitam uma análise mais minuciosa de cada detalhe.

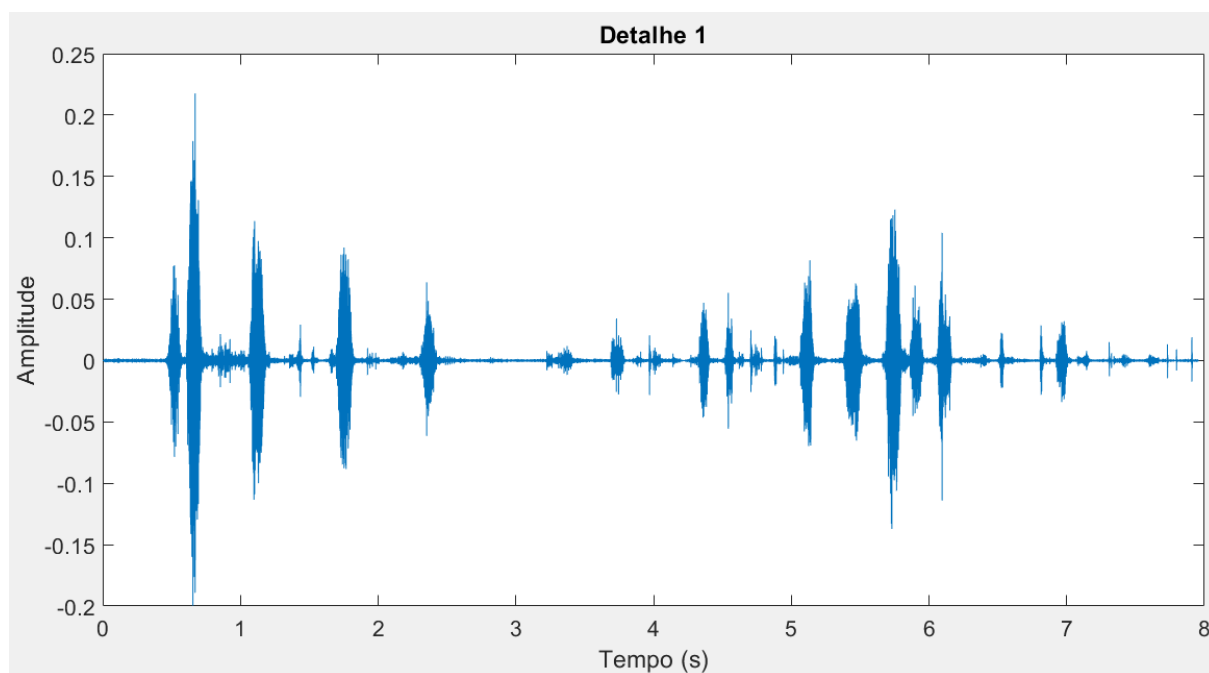


Figura 4. 8 - Detalhe 1 do sinal sonoro da voz.

Conforme é possível perceber, a figura 4.8 permite a análise da faixa de frequência de 11025 Hz a 22050 Hz. Assim, pode-se obter características nos domínios do tempo e da frequência dentro dessa faixa, o que aumenta a precisão dessas informações extraídas.

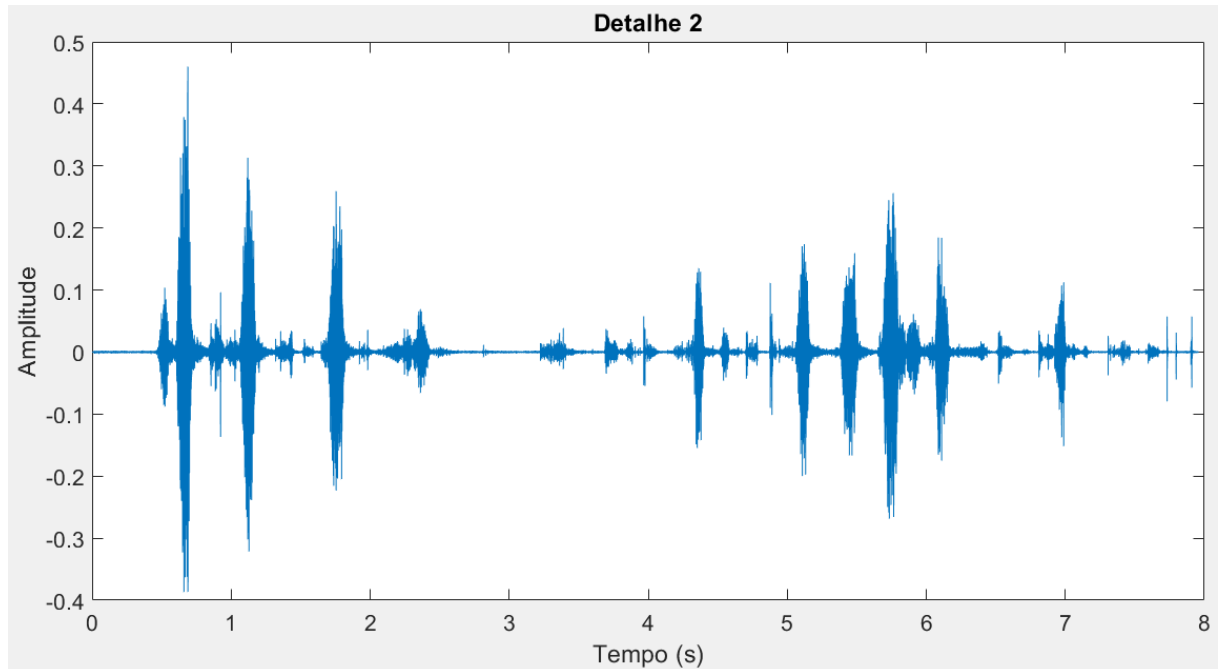


Figura 4. 9 - Detalhe 2 do sinal sonoro da voz.

De acordo com a figura 4.9 e a tabela 4.4, o detalhe 2 do sinal sonoro da voz permite o estudo da faixa de frequência de 5512,50 Hz a 11025 Hz.

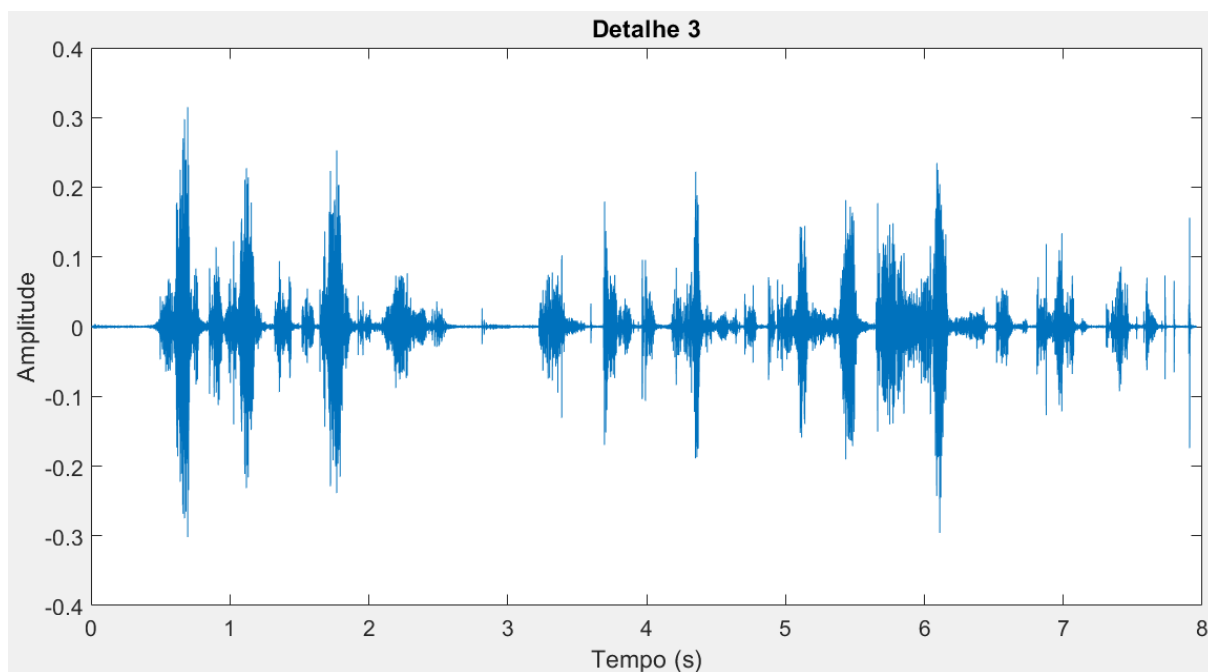


Figura 4.10 - Detalhe 3 do sinal sonoro da voz.

A figura 4.10 mostra o detalhe 3 do sinal sonoro da voz e possibilita o estudo minucioso da faixa de frequência de 2756,25 Hz a 5512,50 Hz.

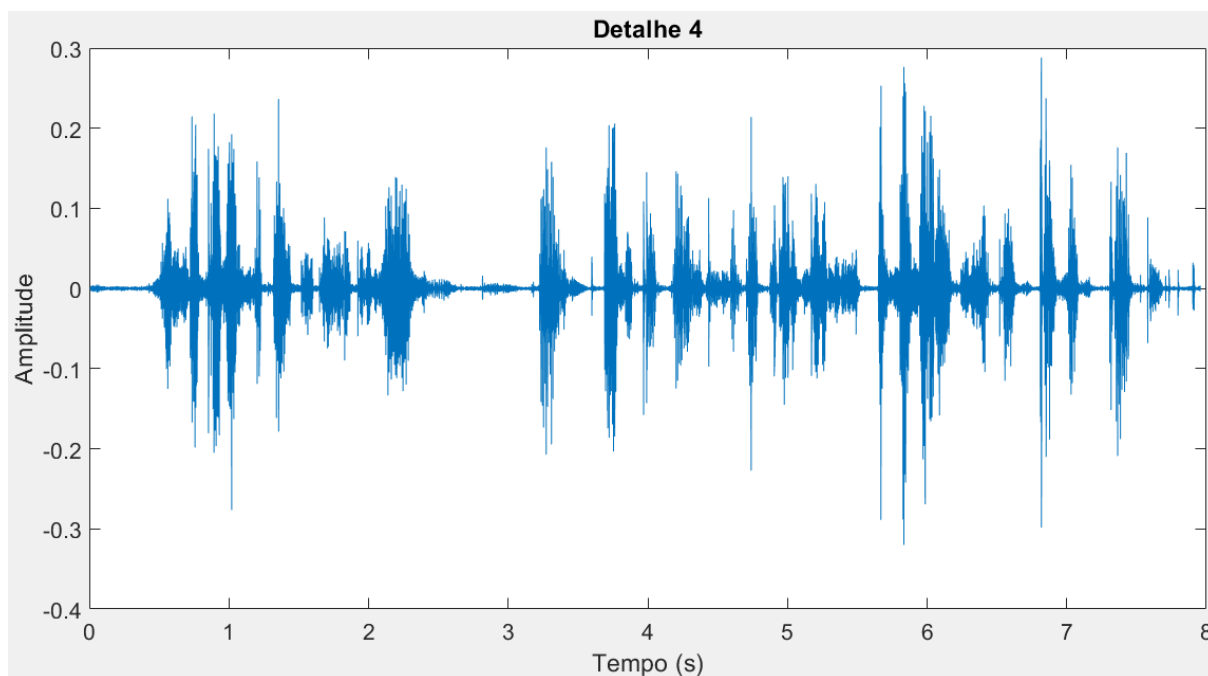


Figura 4.11 - Detalhe 4 do sinal sonoro da voz.

Por fim, a figura 4.11 permite mostra a faixa de frequência de 1378,13 Hz a 2756,25 Hz, o que possibilita a obtenção de características nos domínios do tempo e da frequência dentro dessa faixa.

Como em uma análise multirresolução do sinal nem todas as bandas de frequências são importantes para a extração de características de um sinal, realiza-se, primeiramente, uma análise energética de cada detalhe. Nesse estudo, os detalhes com valores mais elevados de energia serão selecionados e estudados de forma mais aprofundada.

4.5.2 Dados Estatísticos

Após a decomposição do sinal em diversas faixas de frequência, a primeira tentativa de encontrar um padrão que diferencie indivíduos com IC e pessoas saudáveis foi por meio da análise estatística de cada detalhe. A seguir, apresenta-se os dados estatísticos utilizados para o estudo.

A. Média: É definida como o ponto de equilíbrio do sinal e é calculada como se mostra:

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} \quad (4.1)$$

Onde x_i representa os pontos e n o número de elementos do sinal.

B. Mediana: É o valor central de uma série ordenada de números. O valor mediano pode ser encontrado por meio da equação 4.2.

$$M = \frac{n+1}{2} \quad (4.2)$$

C. RMS: É um parâmetro definido como o valor quadrático médio de uma série de números. equação 4.3 fornece o valor RMS de um sinal.

$$X_{RMS} = \sqrt{\frac{\sum_{i=1}^n X_i^2}{n}} \quad (4.3)$$

D. Variância e Desvio Padrão: A variância é baseada no desvio médio, enquanto o desvio padrão é uma medida de dispersão definida como sendo a raiz quadrada da variância. Esses dois parâmetros estatísticos são determinados por meio das equações 4.4 e 4.5, respectivamente.

$$V = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} \quad (4.4)$$

$$\sigma = \sqrt{V} \quad (4.5)$$

E. Amplitude: É a diferença entre o maior e o menor valores de um sinal.

$$A = Max - Min \quad (4.6)$$

F. Segundo Coeficiente de Pearson: É um coeficiente de assimetria que correlaciona a média e a mediana com o desvio padrão. Esse coeficiente, na teoria, varia de -3 a + 3, mas, na prática, dificilmente ultrapassa os valores de -1 e 1. Tal coeficiente é determinado de acordo com a equação 4.7.

$$CP_2 = \frac{3(\bar{X} - M)}{\sigma} \quad (4.7)$$

G. Coeficiente de Variação: O coeficiente de variação possibilita a análise da dispersão do sinal em relação à sua média. Tal parâmetro é encontrado por meio da equação 4.8.

$$CV = \frac{\sigma}{\bar{X}} \quad (4.8)$$

Tais valores estatísticos foram estudados e analisados a fim de identificar um padrão característico que diferencie os indivíduos doentes das pessoas saudáveis.

4.5.3 Transformada Rápida de Fourier (FFT)

Apesar da análise estatística do sinal fornecer parâmetros importantes, eles não são suficientes para identificação das alterações causadas pela insuficiência cardíaca nas vozes dos

indivíduos. Dessa forma, além da análise dos dados no domínio do tempo, foi também realizada a análise no domínio da frequência de cada faixa de frequência dos sinais. Nessas transformadas, os picos mais relevantes para a análise do sinal serão destacados e analisados quanto à relevância para o desenvolvimento das redes neurais.

4.5.4 Análise Mel-Cepstral

A aplicação da análise dos MFCCs tem ganhado relevância por apresentar resultados expressivos na identificação de indivíduos, como mostrado no trabalho de Ittichaichareon et. al. (2012), e na identificação de patologias, como demonstra a pesquisa desenvolvida por Ribeiro et. al. (2014). Nesse sentido, os MFCCs são utilizados na análise da voz humana por sua capacidade de identificar como o aparelho humano interpreta as diferentes frequências. Além disso, essa técnica gera uma grande quantidade de dados que possibilita sua análise por diversas técnicas, como a análise estatística desses dados.

Assim, essa análise será realizada com o objetivo de investigar os coeficientes da escala de Mel, na qual as características globais de cada frame do sinal serão estudadas. Tais parâmetros referem-se à amplitude; ao coeficiente de variação; ao desvio padrão; à média; à mediana; ao valor RMS; à variância; e ao segundo coeficiente de Pearson.

Para o cálculo do MFCC, o tamanho das janelas aplicadas aos áudios é calculado por meio da equação 4.9. Onde, W é o tamanho da janela e F_s é a frequência de amostragem do sinal. Além disso, cabe salientar que o tamanho da janela será o valor inteiro mais próximo de W .

$$W = 0.03 F_s \quad (4.9)$$

Ademais, a função MFCC caracteriza-se por retornar uma matriz de 14 colunas na qual a primeira coluna apresenta os valores da energia logarítmica de cada janela seguido por 13 colunas de coeficientes cepstrais.

4.6 DESENVOLVIMENTO DA RNA

Por fim, os dados gerados pelas análises são extraídos e selecionados de modo a gerar um vetor de características que é utilizado para desenvolver uma RNA capaz de identificar as alterações ocasionadas pela insuficiência cardíaca no ser humano.

A classificação de um parâmetro (média ou frequência fundamental, por exemplo) como relevante se deu por meio da comparação entre os valores obtidos para as pessoas saudáveis e com IC. Assim, o conjunto de parâmetros selecionados formará o vetor de características.

Para o desenvolvimento da RNA a ser utilizada no sistema de identificação de insuficiência cardíaca, o número de neurônios na camada de entrada dependerá diretamente da quantidade de parâmetros presentes no vetor de características. Assim, caso o vetor de características apresente 10 parâmetros, por exemplo, a RNA apresentará 10 neurônios na camada de entrada.

Após a determinação da primeira camada da RNA, deve-se estabelecer a quantidade de neurônios na camada oculta da rede. Nesse sentido, cabe destacar que, na prática, o aumento irrestrito do número de neurônios não acarreta, necessariamente, no ganho de eficiência da rede.

Assim, o número de neurônios da camada oculta se deu por tentativa e erro, visto que não há uma forma padrão para determinar tal valor. Dessa forma, a camada oculta foi determinada por meio da variação da quantidade de neurônios, que começou com 5 neurônios e terminou 50, passando de 5 em 5. Por fim, cada variação na arquitetura da rede foi testada e verificada seu grau de acerto.

Por último, a camada de saída da rede vai ser determinada pela quantidade de classes a ser classificada. Assim, como o presente estudo se limita à classificação de um indivíduo como sendo saudável ou com IC e, portanto, o número de neurônios da última camada da rede será 2.

Além das quantidades de neurônios em cada camada da RNA, algumas características da rede também são importantes para o seu comportamento. Nesse sentido, as funções de transferência têm papel a ser destacado.

Para a primeira camada, foi utilizada a função de transferência linear, de modo que o parâmetro de entrada não é alterado por essa função. Dessa forma, a primeira camada da RNA servirá apenas como um meio transmissor do valor de entrada para a segunda camada.

Já na camada oculta, optou-se pela utilização da função de transferência tangente hiperbólica por sua característica de lidar com valores negativos, uma vez que os sinais sonoros das vozes tendem a apresentar valores diversos e imprevisíveis.

Por fim, a função softmax foi aplicada na camada de saída da rede, de modo a classificar, probabilisticamente, a classe do sinal. Nesse sentido, espera-se que a rede, além de classificar o sinal em “indivíduo saudável” ou “indivíduo com IC”, apresente a probabilidade de certeza do acerto.

No que concerne o processo de aprendizagem da rede, foi utilizado algoritmo de aprendizagem de retropropagação do erro. Tal algoritmo ajusta os pesos associados ao sinal de entrada continuamente de forma a atingir o menor erro possível. Assim, a rede funcionará como um sistema fechado que reavalia o erro e o ajusta até que este não tenha mais mudança significativa. Ao atingir tal ponto, diz-se que a rede está treinada.

Após o desenvolvimento das RNAs para as vozes masculina e feminina, foi iniciado o processo de elaboração da interface gráfica de forma a facilitar o seu uso pelos profissionais da saúde. A interface consiste de um aplicativo de computador, no qual o usuário terá acesso ao pressionar duas vezes o seu ícone na área de trabalho. Ademais, o aplicativo fornecerá dados, gráficos e um possível diagnóstico para que o médico avalie e tome as devidas providências.

Portanto, espera-se que esse sistema forneça valores de probabilidade para cada classe analisada (saudável e com IC); assim, o aplicativo desenvolvido vai sugerir um possível diagnóstico ao médico responsável, além de mostrar gráficos e tabelas com valores padrões para pessoas saudáveis e com IC. Conclui-se assim que o sistema não pretende gerar um diagnóstico definitivo, mas apenas auxiliar o médico no processo de classificação da IC.

4.7 CONCLUSÃO

Nesse capítulo, foi apresentada a metodologia para o desenvolvimento desta pesquisa.

Primeiramente, foi elaborado o fluxograma da pesquisa, no qual todos os passos necessários para o desenvolvimento do estudo são expostos.

Para a aquisição dos dados, foi apresentado os locais de gravação assim como a frase repetida pelo participante e o gravador utilizado. Depois, mostra-se os critérios de inclusão e exclusão da pesquisa.

Após a etapa de aquisição, separou-se os dados de acordo com o sexo e idade do indivíduo. Tal procedimento é denominado pré-processamento.

No processamento dos dados, fez-se o uso das seguintes técnicas de análise de sinais: DWT; análise estatística; FFT e MFCC.

Para o desenvolvimento das RNAs, os dados extraídos na etapa anterior são utilizados para a criação das RNAs masculina e feminina. Após esse processo, elabora-se o aplicativo de forma a facilitar o uso do programa pelos profissionais da saúde. O aplicativo consiste em uma interface gráfica que fornecerá dados, gráficos e um possível diagnóstico para que o médico avalie e tome as devidas providências

CAPÍTULO V

RESULTADOS E DISCUSSÕES

O capítulo 5 desta dissertação apresenta os resultados obtidos por meio da aplicação de técnicas de análise de sinais, da utilização de RNA e do desenvolvimento de interface gráfica para a identificação de IC.

Seguindo os critérios de exclusão e inclusão estabelecidos, foram gravadas 68 vozes, sendo 44 homens e 24 mulheres, de pacientes diagnosticados com insuficiência cardíaca pelo INCOR. Além disso, 16 pacientes, 10 homens e 6 mulheres, com insuficiência cardíaca diagnosticada pelo hospital metropolitano de João Pessoa foram coletados. Por outro lado, foram gravadas as vozes de 58 pessoas em ambiente extra hospitalar, sendo 29 homens e 29 mulheres que não apresentavam sintomas da doença e, pelo critério da classe funcional (NYHA), consideradas saudáveis.

5.1 EXTRAÇÃO E SELEÇÃO DAS CARACTERÍSTICAS DAS VOZES MASCULINAS

Nessa sessão apresenta-se os resultados obtidos por meio da aplicação das diferentes técnicas de análise de sinais nos sinais sonoros das vozes masculinas aquistadas.

5.1.1 Análise Multirresolução dos Sinais

A análise em multirresolução consistiu em segmentar os sinais em 11 faixas de frequências. Tal quantidade de detalhes foi escolhida pela viabilidade de realizar um estudo minucioso dos dados.

Como essa análise gera uma grande quantidade de dados, a seleção das bandas relevantes para a pesquisa ajuda a diminuir o custo computacional do sistema elaborado e aumentar a eficiência da rede. Dessa forma, a análise energética dos sinais possibilita uma primeira diminuição na quantidade de dados a serem estudados para a identificação da IC por meio da voz. Nesse sentido, a figura 5.1 apresenta a evolução da energia para cada detalhe dos sinais, enquanto a figura 5.2 mostra a seleção dos detalhes mais energéticos.

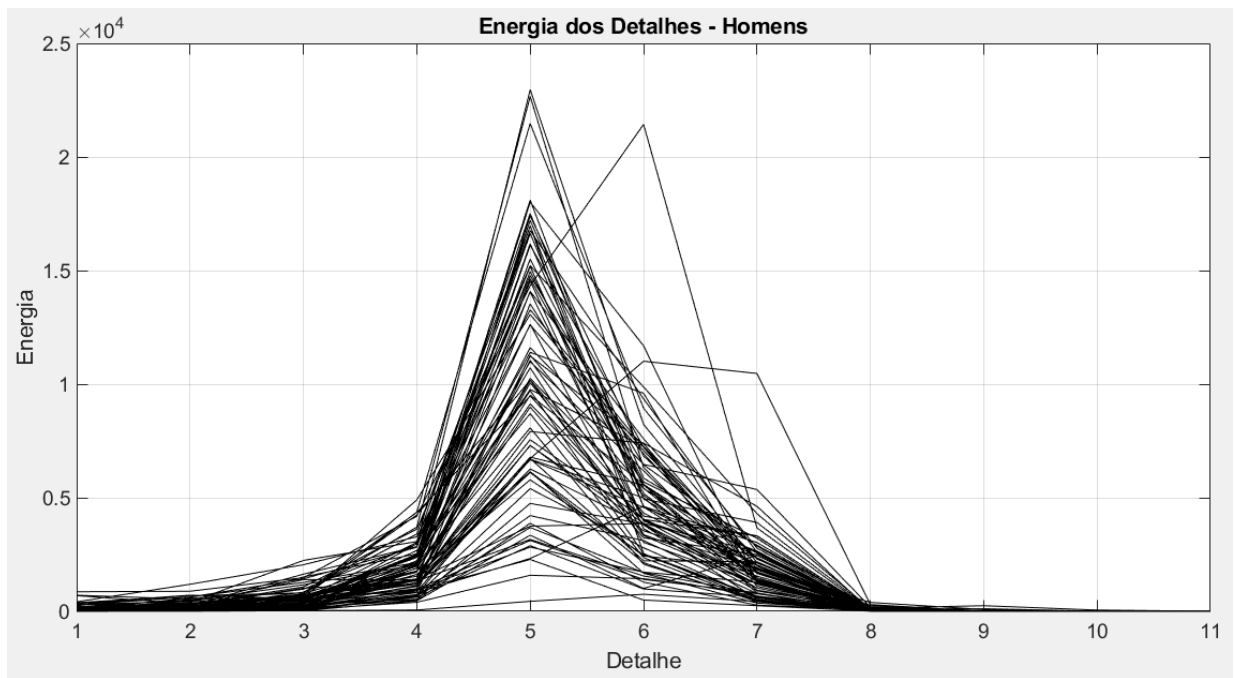


Figura 5.1 - Evolução energética dos detalhes – Masculino.

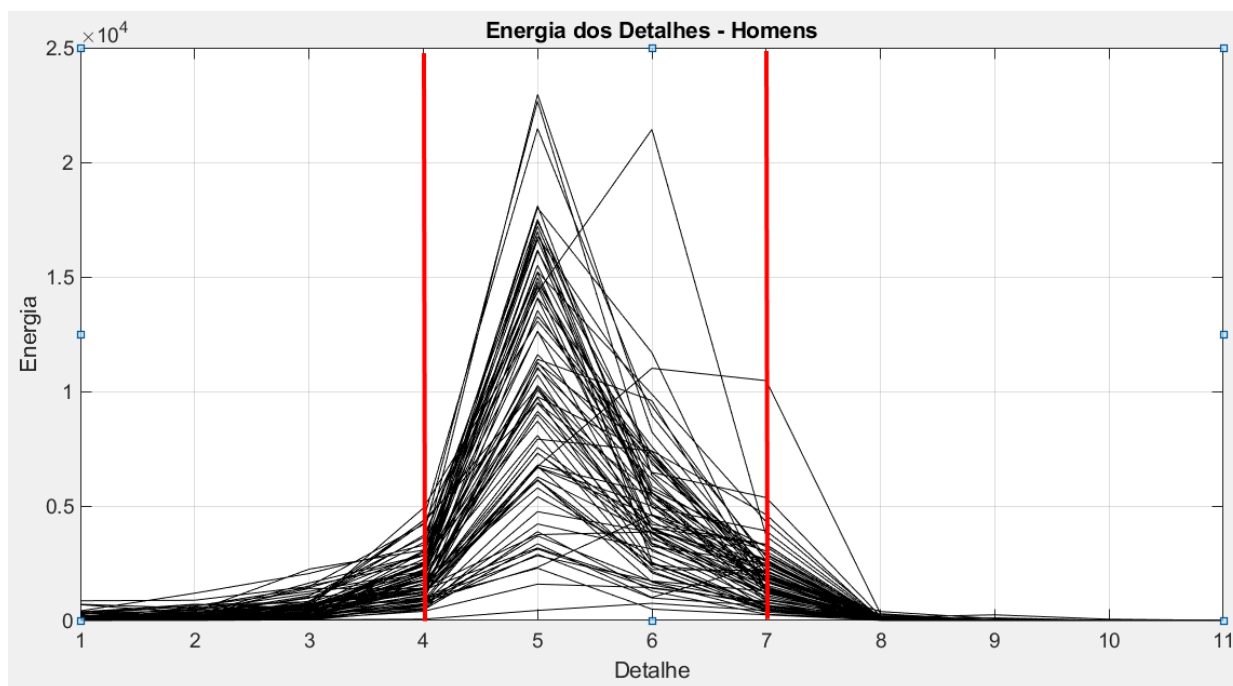


Figura 5.2 - Seleção dos detalhes – Masculino.

Ao observar a figura 5.2, infere-se que os detalhes 4, 5, 6 e 7 se sobressaem em relação aos demais. De acordo com a teoria já estabelecida, as cordas vocais de um ser humano do sexo masculino vibram em torno de 125 vezes por segundo (125 Hz), o que se aproxima do detalhe 7 (172,27 – 344,53 Hz). Além disso, ruídos e distorções no trato vocal alteram o padrão da voz. Portanto, os detalhes 4, 5, 6 e 7 serão objetos de estudo para que um padrão seja identificado de modo a diferenciar uma pessoa saudável de um indivíduo com IC.

Assim, as figuras 5.3 a 5.6 apresentam os detalhes 4, 5, 6 e 7, respectivamente, de uma decomposição em 11 faixas de frequência de um sinal sonoro de uma voz masculina.

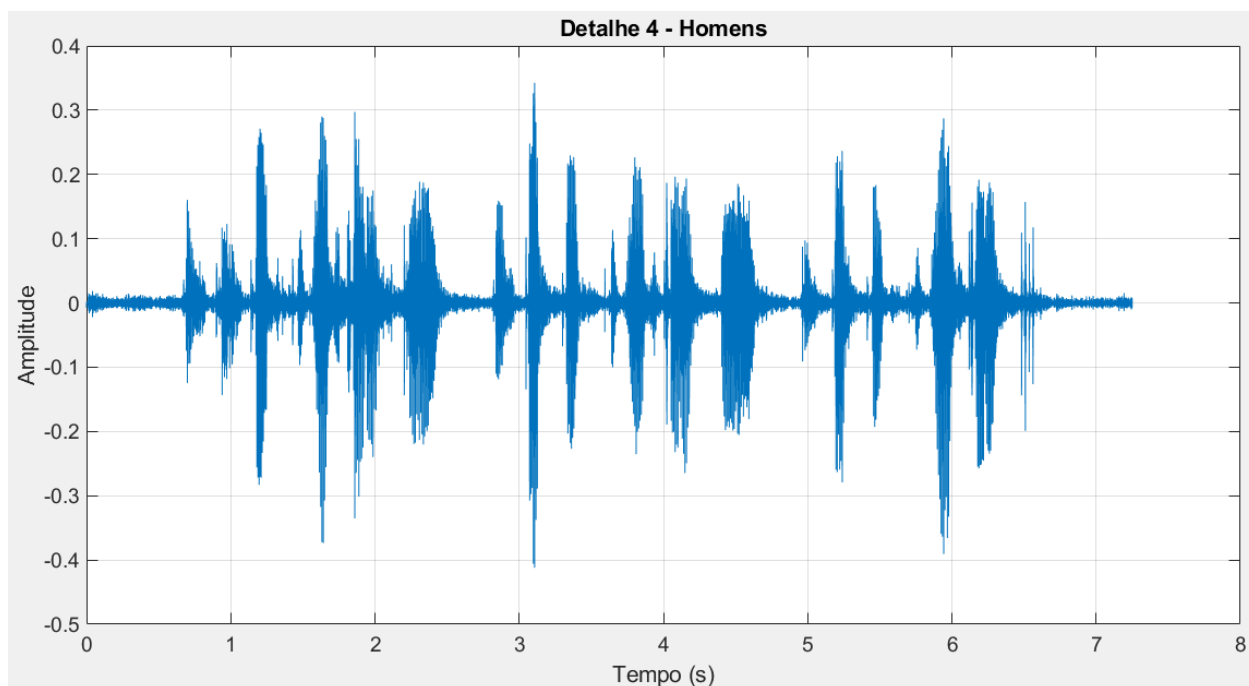


Figura 5.3 - Detalhe 4 – Masculino.

A figura 5.3 apresenta o sinal sonoro de uma voz masculina limitado ao detalhe 4. Dessa forma, ao aplicar as técnicas de análise de sinais, os parâmetros extraídos serão referentes à faixa de frequência que vai de 1378,13 Hz a 2756,25 Hz, o que possibilita uma análise mais minuciosa dessa faixa.

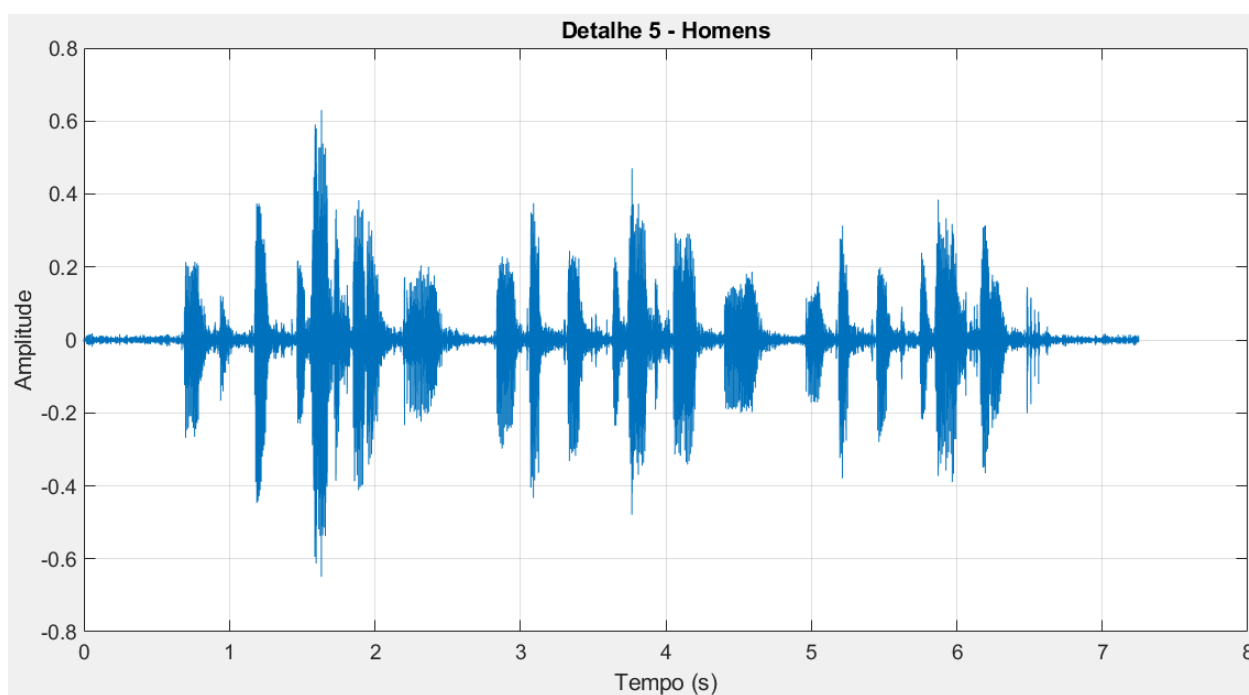


Figura 5.4 - Detalhe 5 – Masculino.

A figura 5.4 apresenta o sinal sonoro da mesma voz representada na figura 5.3; porém esta mostra o detalhe 5 do sinal, o que representa a faixa de frequência que vai de 689,06 a 1378,13 Hz.

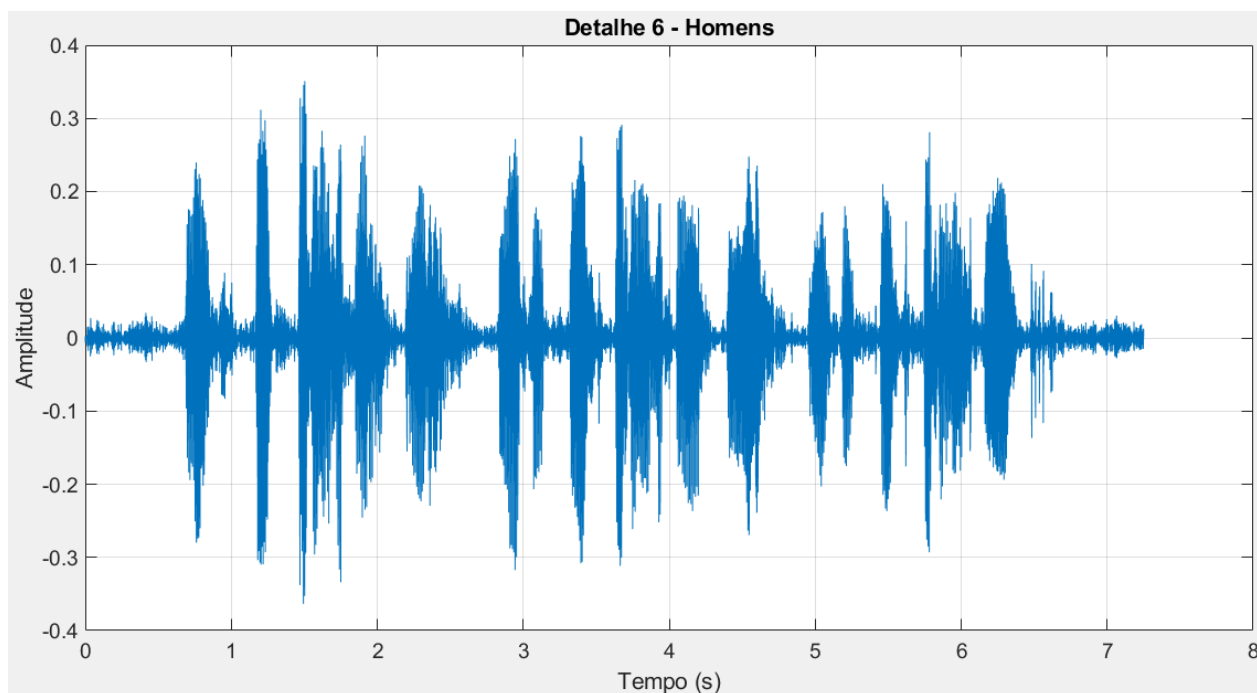


Figura 5.5 - Detalhe 6 – Masculino.

Para o detalhe 6, a figura 5.5 mostra o sinal sonoro da voz masculina de um indivíduo saudável compreendido na faixa de frequência que vai de 344,53 a 689,06 Hz. Assim, é possível extrair e selecionar parâmetros específicos desse detalhe.

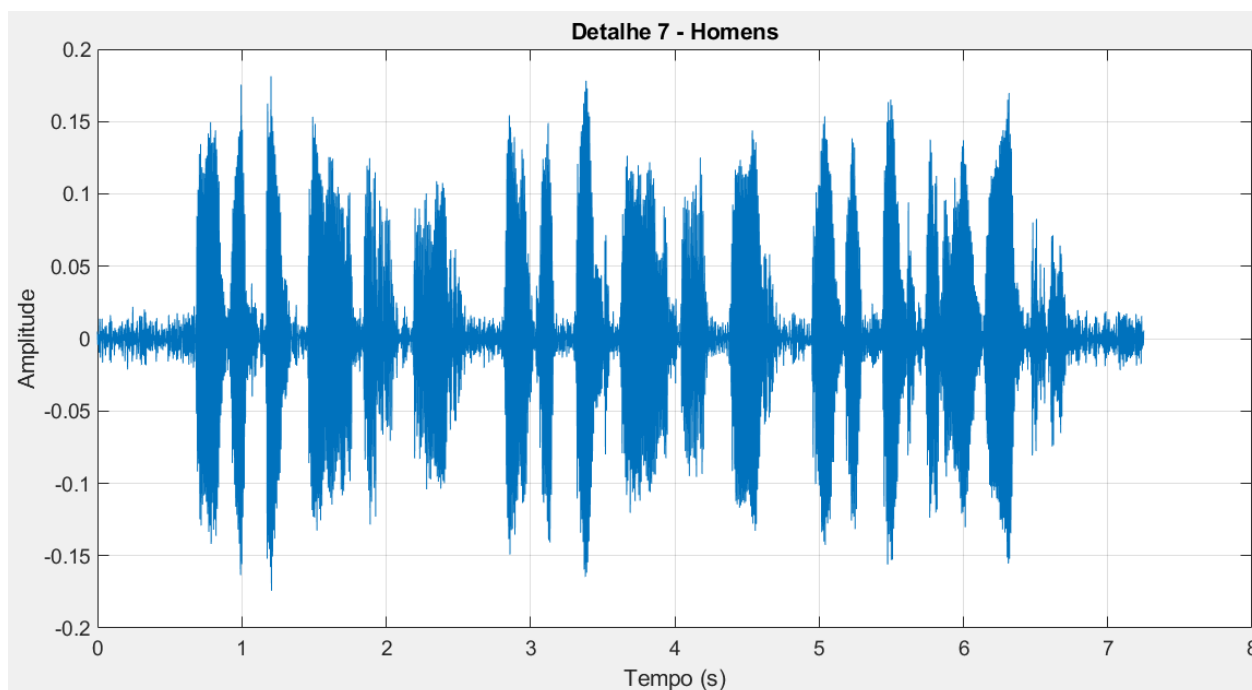


Figura 5.6 - Detalhe 7 – Masculino.

Por fim, o detalhe 7, representado na figura 5.6, apresenta o sinal delimitado entre 172,27 e 344,53 Hz. Assim, uma análise mais detalhada dessa faixa de frequência pode ser realizada, visto que as demais frequências não interferem no estudo.

5.1.2 Análise Estatística

Para cada um dos detalhes selecionados na seção anterior uma análise estatística foi desenvolvida. Assim, as figuras 5.7 a 5.14 diferenciam, respectivamente, os valores de média, mediana, variância, desvio padrão, RMS, amplitude, segundo coeficiente de Pearson e coeficiente de variação entre os homens saudáveis e os com insuficiência cardíaca.

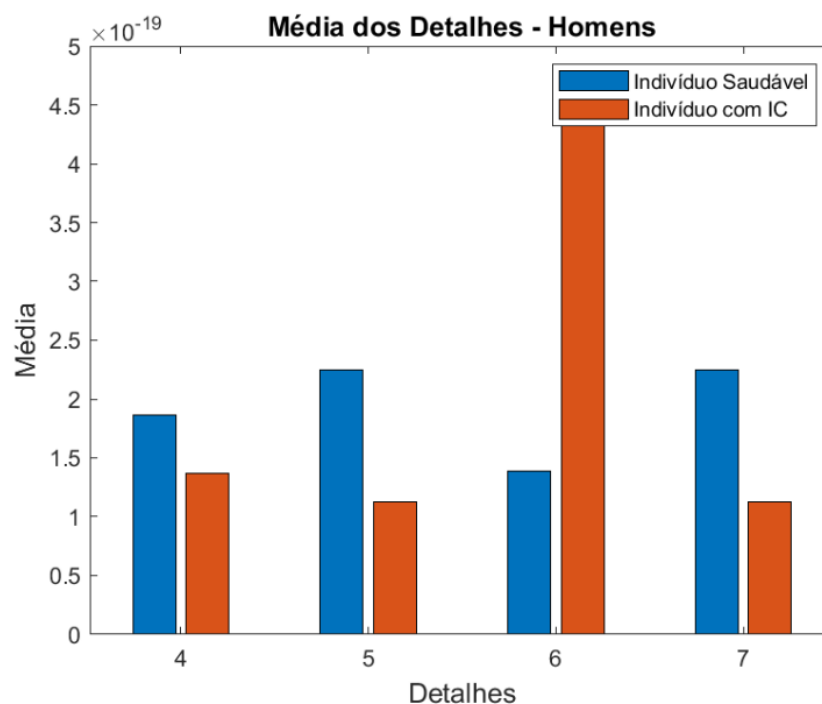


Figura 5.7 - Média dos detalhes – Masculino.

A análise das médias mostra que o detalhe 6 apresenta uma grande diferença entre os valores encontrados para os homens saudáveis e para os homens com IC. Assim, a média no detalhe 6 se apresenta mais elevada nos homens com IC do que nos indivíduos saudáveis. Por outro lado, as médias nos outros detalhes são superiores para os homens saudáveis, porém a variação de cada detalhe individualmente não apresenta um padrão que possibilite distinguir os homens saudáveis dos doentes.

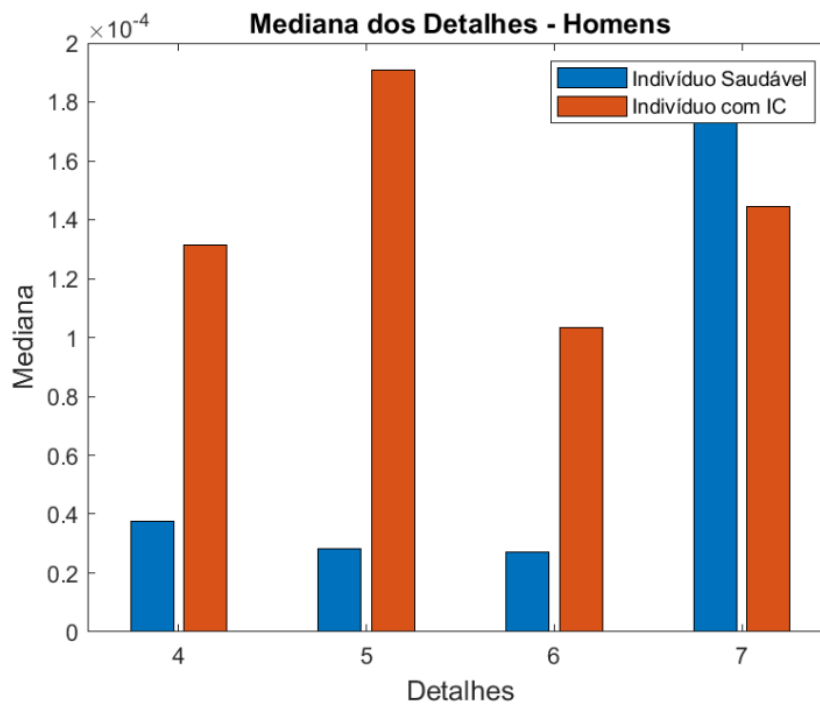


Figura 5.8 - Mediana dos detalhes – Masculino.

Para a mediana, os detalhes 4, 5 e 6 apresentam uma diferença relevante entre os indivíduos saudáveis e com IC. A análise dos detalhes destaca que, para o homem com IC, a mediana nessas faixas de frequências é substancialmente maior que a mediana dos homens saudáveis.

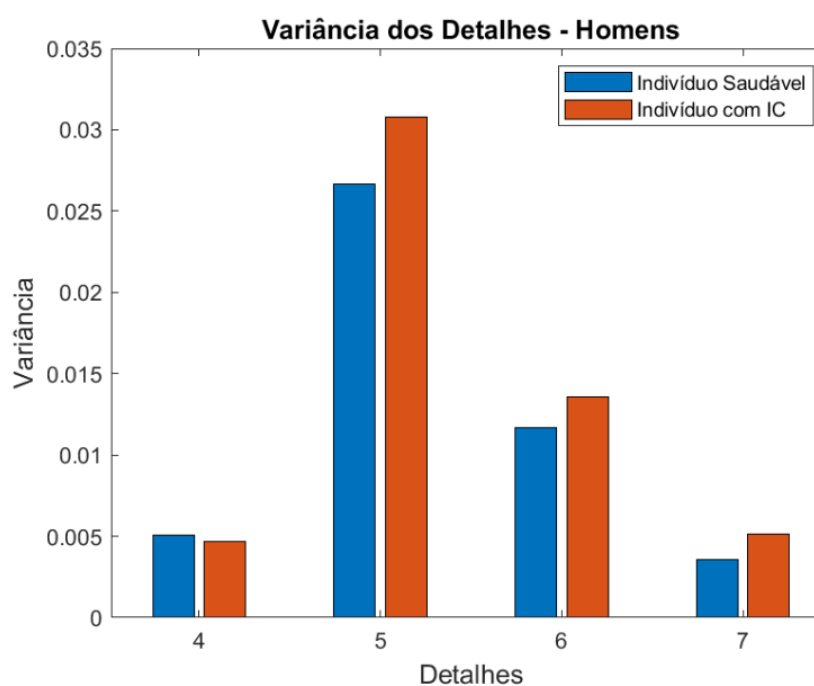


Figura 5.9 - Variância dos detalhes – Masculino.

A figura 5.9 apresenta os valores relativos à variância dos dados obtidos. Nessa figura, observa-se que a variância apresenta uma pequena variação nos detalhes 5,6 e 7, porém essa diferença não é significativa para a classificação das vozes masculinas saudáveis e com IC.

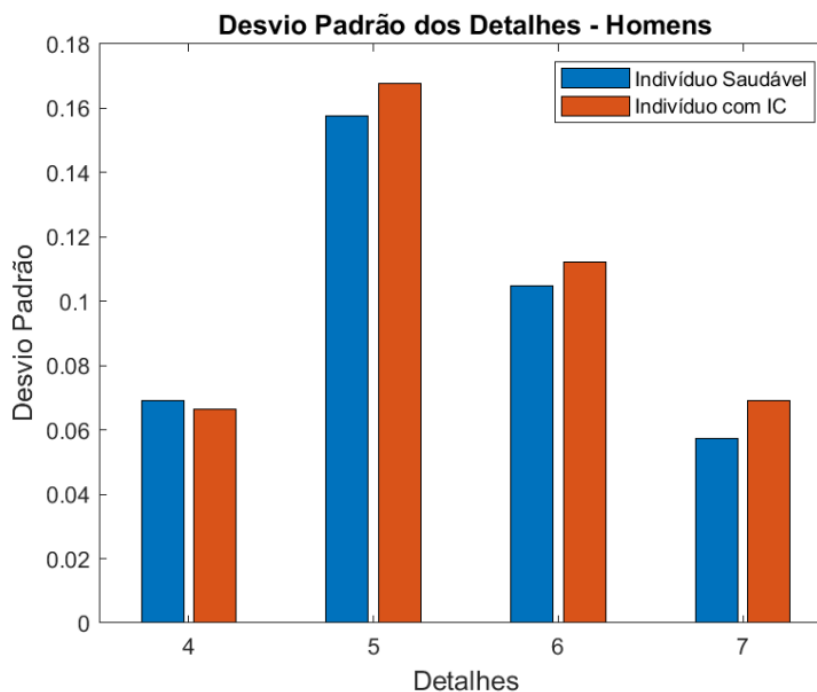


Figura 5.10 - Desvio Padrão dos detalhes – Masculino.

O desvio padrão é definido como sendo a raiz quadrada da variância. Assim, espera-se que o padrão observado para a variância se repita para o desvio padrão. Nesse sentido, a figura 5.10 mostra que os detalhes 4, 5, 6 e 7 apresentam uma pequena diferença entre os desvios para o homem saudável e para o homem com IC. Dessa forma, optou-se pela não aplicação do desvio padrão na identificação da patologia.

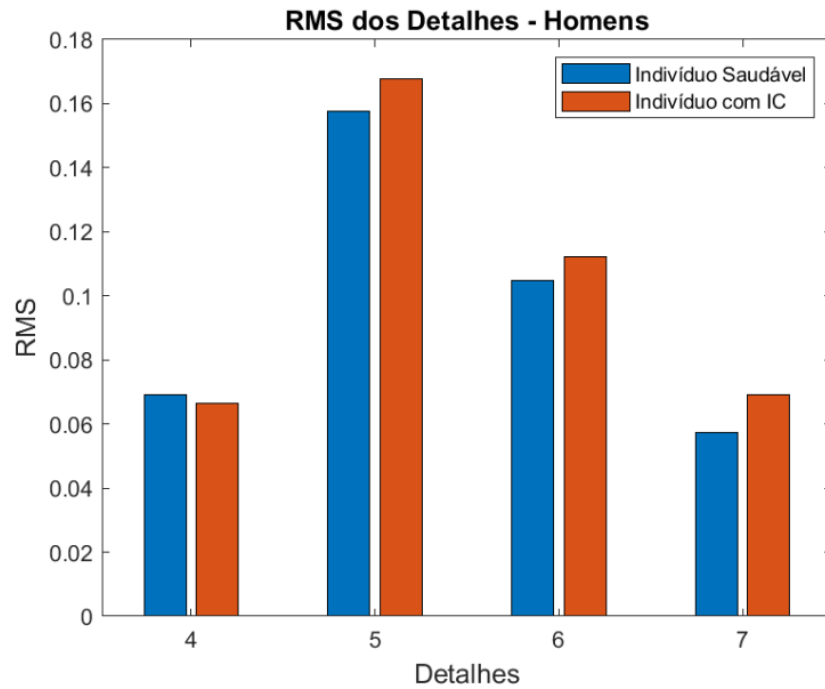


Figura 5.11 - RMS dos detalhes – Masculino.

Para o valor RMS, a figura 5.11 mostra esse parâmetro pouco se altera entre os indivíduos saudáveis e com IC, independentemente do detalhe analisado.

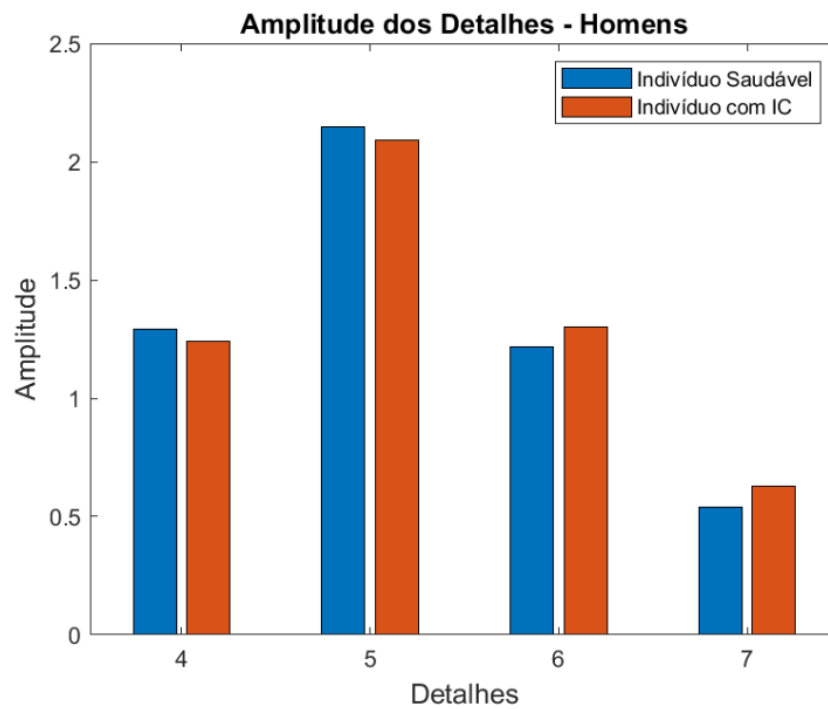


Figura 5. 12 - Amplitude dos detalhes – Masculino.

A amplitude dos sinais, em cada detalhe, é mais um parâmetro que pouco se altera ao se analisar indivíduos saudáveis ou doentes.

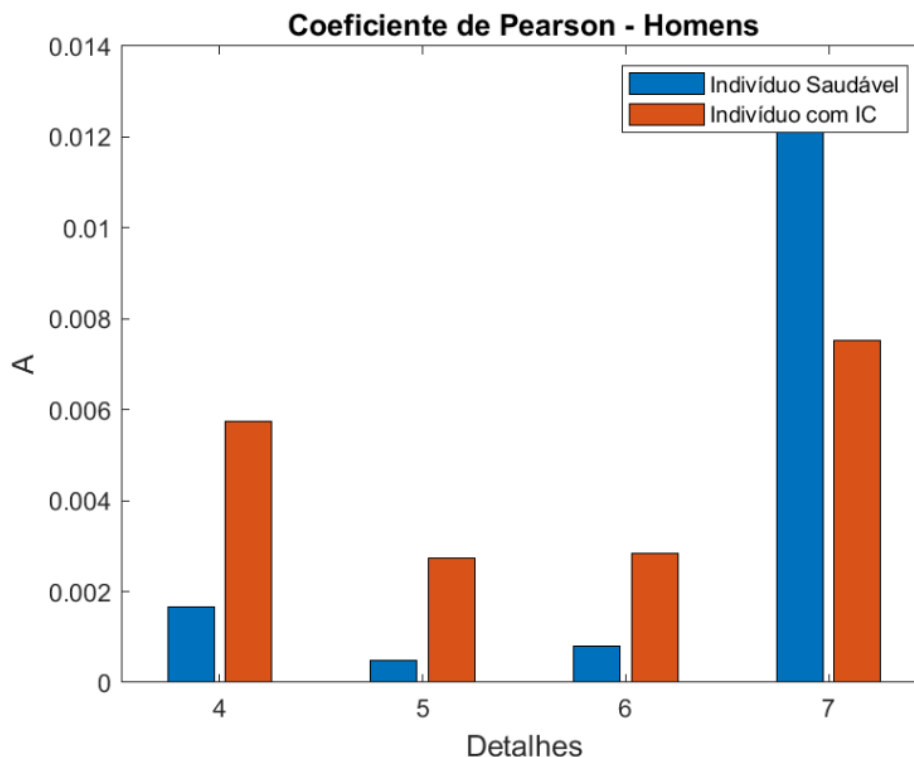


Figura 5. 13 - Coeficiente de Pearson dos detalhes – Masculino.

O coeficiente de Pearson é um parâmetro que correlaciona a média e a mediana com o desvio padrão, medindo a assimetria de um sinal. Assim, a figura 5.13 mostra que nos detalhes 4, 5 e 6, as vozes dos homens com insuficiência cardíaca são mais assimétricas que a dos homens saudáveis. Por outro lado, no detalhe 7, a situação se inverte, onde a voz masculina saudável é mais assimétrica que a voz de um doente.

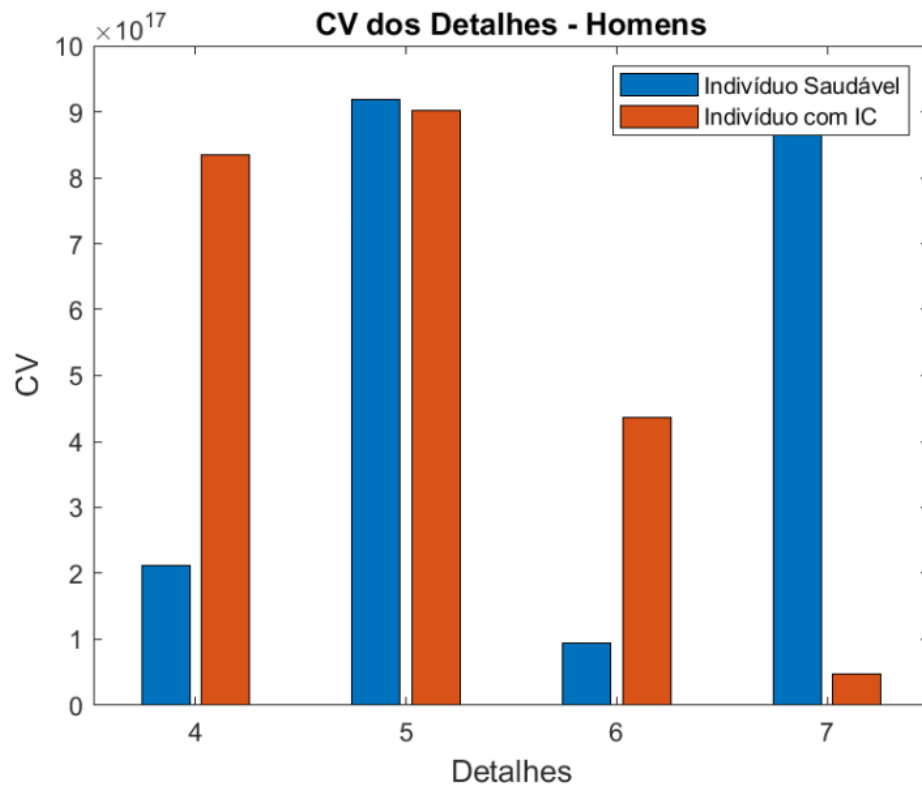


Figura 5.14 - Coeficiente de Variação dos detalhes – Masculino.

Por fim, a figura 5.14 mostra o comportamento do coeficiente de variação dos sinais em análise. Nesse sentido, o coeficiente de variação no detalhe 4 apresenta uma significativa diferença entre os valores obtidos, onde o coeficiente para os homens saudáveis é menor que o observado para os homens com IC. Já para o detalhe 7, o coeficiente de variação se mostra bem mais alto para as vozes masculinas saudáveis.

Assim, ao analisar os dados estatísticos referentes aos detalhes dos sinais das vozes masculinas, foi possível identificar algumas diferenças entre os sinais provenientes dos indivíduos saudáveis e os sinais dos pacientes com insuficiência cardíaca. Essas informações extraídas serão aplicadas, em conjunto as demais, para o desenvolvimento e validação da RNA masculina.

De modo a apresentar os dados relevantes para o desenvolvimento da RNA mais didaticamente, a tabela 5.1 assinala tais características.

Detalhe	4	5	6	7
Média	-	-	X	-
Mediana	X	X	X	-
Variância	-	-	-	-
Desvio Padrão	-	-	-	-
RMS	-	-	-	-
Amplitude	-	-	-	-

Coeficiente de Pearson	X	X	X	X
Coeficiente de Variação	X	X	-	X

Tabela 5. 1 - Características relevantes para o desenvolvimento da RNA – Masculino.

Portanto, a tabela 5.1 sumariza todo o processo de seleção dos parâmetros relevantes para o desenvolvimento do sistema de auxílio ao médico para o diagnóstico da IC, onde os 11 parâmetros selecionados por meio da análise estatística dos sinais estão marcados com X.

5.1.3 Análise no Domínio da Frequência

Nesse trabalho, a análise no domínio da frequência do sinal ocorre por meio da aplicação da FFT em cada detalhe selecionado para o estudo. Dessa forma, tal estudo consiste na seleção do pico predominante e sua respectiva frequência. Além disso, a energia de cada detalhe foi levada em consideração nessa fase da análise do sinal. As figuras 5.15 a 5.18 mostram o comparativo das assinaturas espectrais, que é o padrão de um sinal no domínio da frequência, entre pessoas saudáveis e doentes nos detalhes 4, 5, 6 e 7, respectivamente.

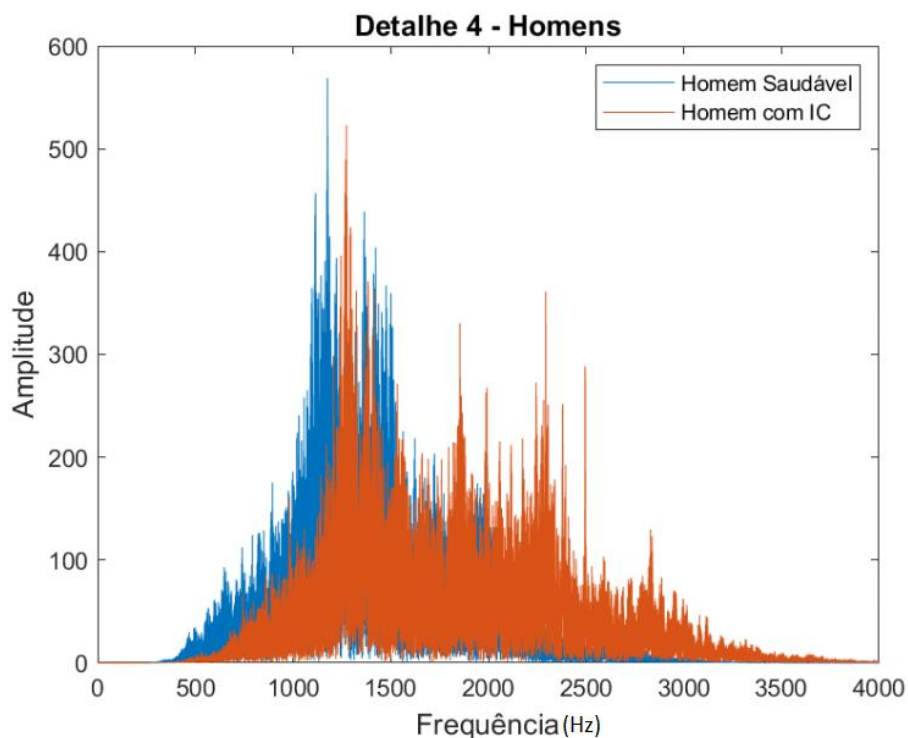


Figura 5.15 - FFT do detalhe 4 – Masculino.

A figura 5.15 apresenta o padrão do sinal sonoro de uma voz masculina no domínio da frequência e limitado ao detalhe 4. Dessa forma, a assinatura espectral desses sinais está restrita à faixa de frequência que de 1378,13 Hz a 2756,25 Hz. Assim, a análise do detalhe 4 para os homens saudáveis e com IC não apresentam diferenças significativas entre os espectros analisados, visto que eles quase sobrepõem um ao outro, a ponto de determinar um padrão. Portanto, o detalhe 4 não foi utilizado para o desenvolvimento da RNA masculina.

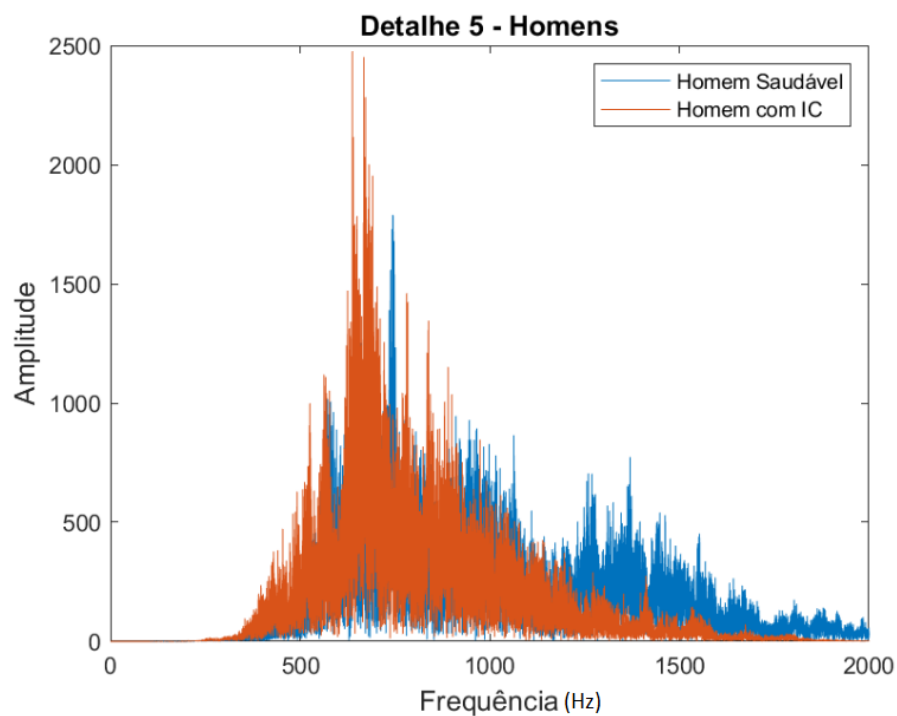


Figura 5. 16 - FFT do detalhe 5 – Masculino.

Para a faixa de frequência que vai de 689,06 a 1378,13 Hz, representado pelo detalhe 5 e mostrado na figura 5.16, percebe-se que há um pico em destaque, no qual a amplitude do sinal referente ao homem com IC apresenta uma amplitude significativamente maior que a amplitude relativa ao homem saudável. Dessa forma, o conjunto pico/frequência é selecionado para compor o vetor de características que será utilizado no desenvolvimento da RNA.

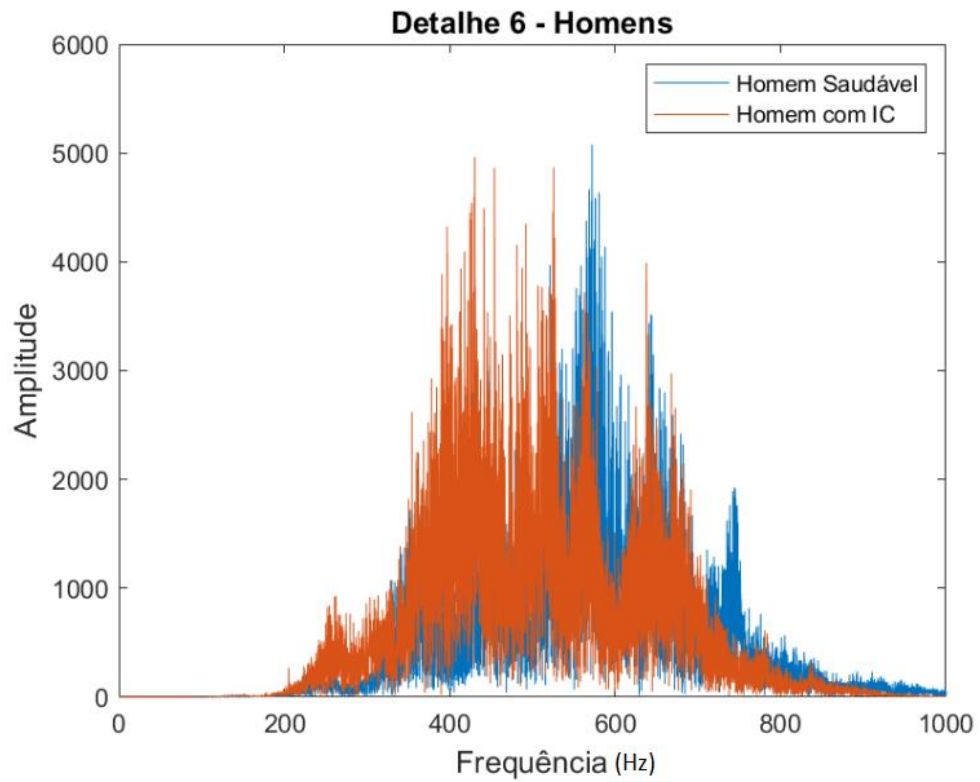


Figura 5.17 - FFT do detalhe 6– Masculino.

Para o detalhe 6, delimitado na faixa de frequência 344,53 – 689,06 e mostrado na figura 5.17, a análise das assinaturas espectrais não levou a identificação de um padrão diferenciável para a classificação da doença.

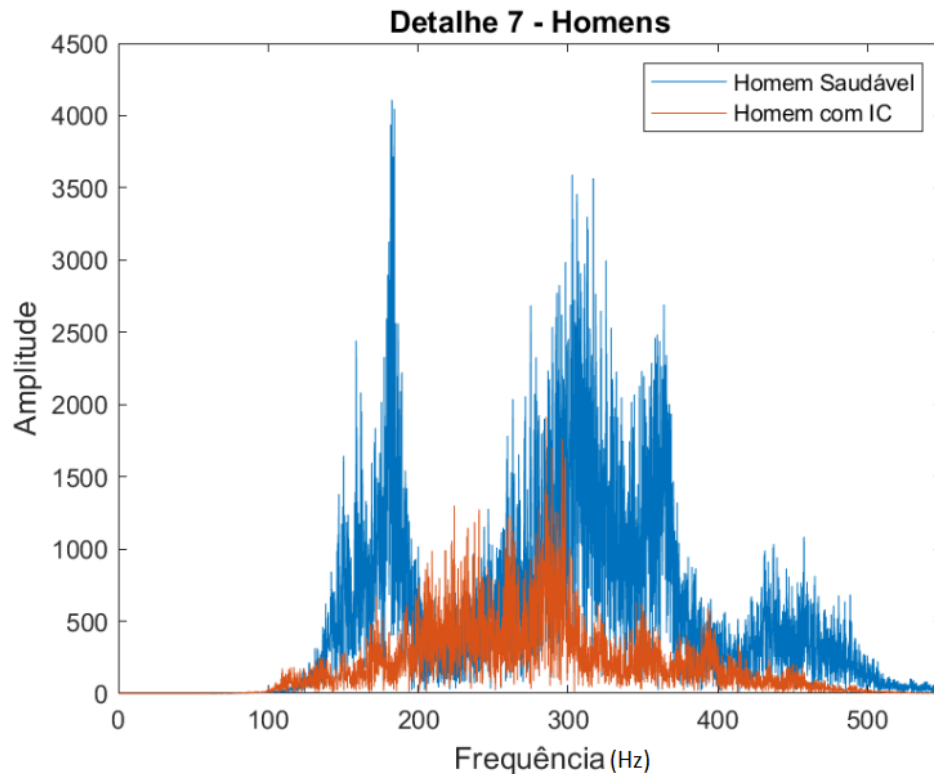


Figura 5.18 - FFT do detalhe 7 – Masculino.

A assinatura espectral do detalhe 7, mostrado na figura 5.18, apresenta um padrão de sinal que está contido na faixa de frequência que vai de 172,27 a 344,53 Hz. Ao analisar esse detalhe, percebe-se que há dois picos proeminentes relacionados ao homem saudável. Dessa forma, para o detalhe 7, seleciona-se o valor de amplitude desses picos e suas respectivas frequências, de modo a utilizá-los para a elaboração da RNA.

Dessa forma, as figuras 5.15 a 5.18 representam padrões identificados por meio da análise de vários sinais estudados. Tais dados foram examinados comparando homens dentro de uma mesma classe de idade.

Ao realizar a análise espectral dos detalhes, foi possível identificar padrões diferentes entre os indivíduos saudáveis e os indivíduos com insuficiência cardíaca nos detalhes 5 e 7.

Além disso, o nível energético de cada faixa de frequência também foi examinado. Nesse sentido, a figura 5.19 expõe a média da energia em cada banda.

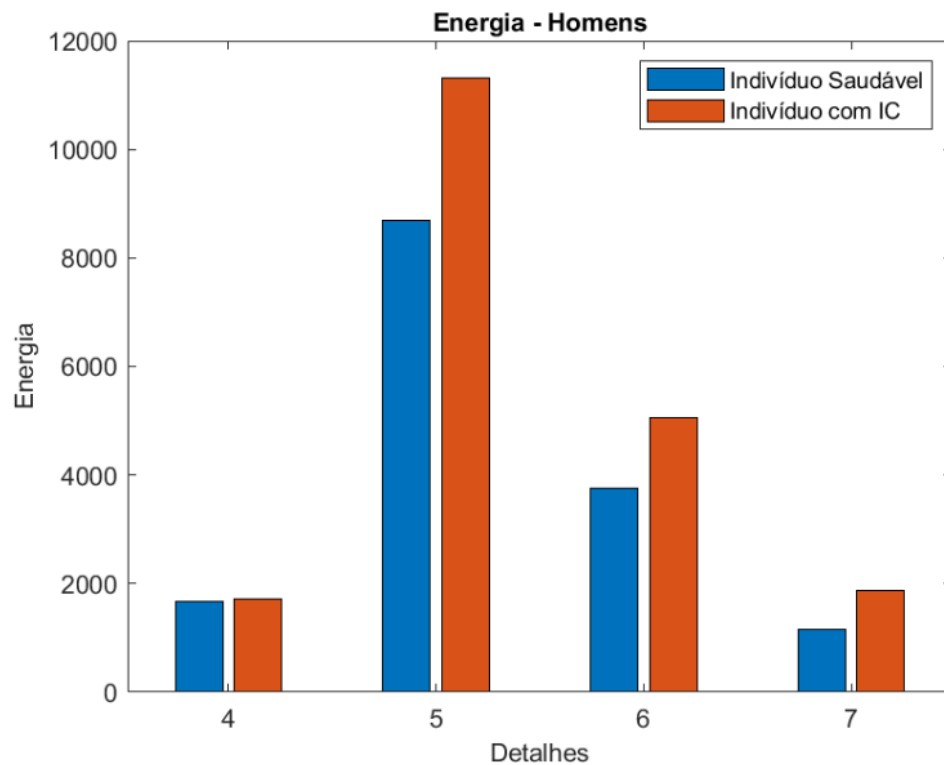


Figura 5.19 - Energia dos detalhes – Masculino.

Assim, por meio da figura 5.19, observa-se que os detalhes 4 e 7 não apresentam grandes diferenças entre os indivíduos saudáveis e com IC. Por outro lado, os detalhes 5 e 6 se mostram relevantes, pois as vozes dos homens doentes apresentam, em média, uma maior energia associada a essas faixas. Assim, pode-se dizer que a energia associada a esses detalhes se mostra importante para a determinação do padrão das vozes masculinas com IC (maior energia) e saudáveis (menor energia).

Por fim, a tabela 5.2 apresenta as características importantes extraídas por meio da análise no domínio da frequência dos sinais.

Parâmetro	Detalhe 4	Detalhe 5	Detalhe 6	Detalhe 7
Energia	-	X	X	-
Picos (FFT)	-	X	-	X

Tabela 5. 2 - Análise no domínio da frequência – Masculino.

Portanto, a tabela 5.2 apresenta, de forma resumida, os parâmetros selecionados nessa seção de análise dos sinais no domínio da frequência para o desenvolvimento da RNA de identificação da insuficiência cardíaca. Nessa tabela, os 4 dados relevantes selecionados por meio da análise dos sinais no domínio da frequência estão marcados com X.

5.1.4 Análise Mel-Cepstral

Por fim, com intuito de desenvolver o sistema de identificação das alterações nas vozes masculinas devido à presença da insuficiência cardíaca, foi realizada a análise cepstral do sinal. Essa técnica apresenta um ganho importante para o sistema pois possibilita a aplicação dos coeficientes cepstrais de frequência mel que representam as mudanças do sistema vocal.

Por padrão, a metodologia empregada para obter os MFCCs particiona o sinal em frames de acordo com o seu tamanho, com o comprimento de janela adotado e com o comprimento da sobreposição. Além disso, utiliza-se os valores de 3% e de 2,5% da frequência de amostragem para os valores de tamanho da janela e da sobreposição, respectivamente. Dessa forma, como a taxa de aquisição para todos os sinais obtidos foi de 44,1 KHz, os comprimentos de janela e de sobreposição foram, respectivamente, 1323 e 1103 pontos do sinal.

Aplicando esses valores padrão, a análise cepstral gera 14 colunas de dados, nas quais a primeira se refere ao valor logarítmico da energia do sinal e as outras colunas se referem aos valores dos coeficientes cepstrais do sinal.

Como o uso dos MFCCs gera uma grande quantidade de dados, opta-se pela análise de suas características globais; são elas: amplitude; coeficiente de Pearson; coeficiente de variação (CV); desvio padrão; média; mediana; RMS e variância.

Portanto, as figuras 5.20 a 5.27 trazem uma perspectiva geral de cada característica analisada. É importante mencionar que as figuras mostram apenas um sinal que representa os demais e, dessa forma, a análise pode diferenciar da representação feita pela imagem.

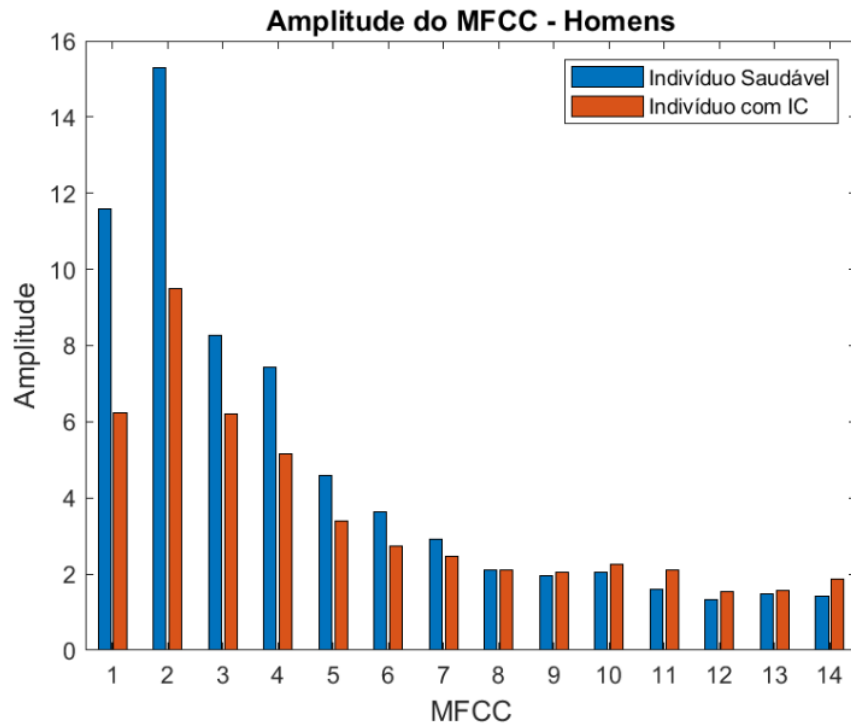


Figura 5.20 - Amplitude dos MFCCs – Masculino.

A figura 5.20 apresenta a amplitude dos MFCCs, os quais são distribuídos em 13 quadros (barras 2 a 14). Além disso, a amplitude da energia logarítmica também é calculada e representada pela primeira barra da figura. Assim, ao comparar os valores obtidos por meio dos sinais mostram que os frames 1 e 3 (barras 2 e 4) são os que apresentam valores significativamente distintos. Quanto à energia (barra 1), vê-se que esta apresenta valores diferentes para os casos em estudo. Assim, a energia e os frames 1 e 3 apresentam um padrão que diferenciam as vozes dos indivíduos saudáveis e com IC.

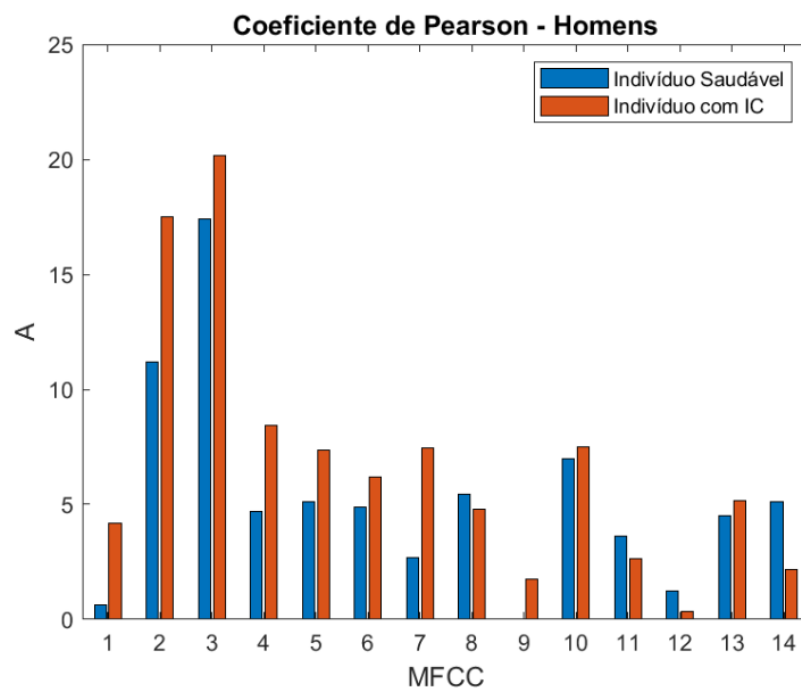


Figura 5. 21 - Coeficiente de Pearson dos MFCCs – Masculino.

Para o segundo coeficiente de Pearson (figura 5.21), os frames 1, 2, 3 e 5 mostram valores significativamente diferentes para homens doentes e saudáveis, onde os indivíduos com IC apresentaram valores mais elevados. Dessa forma, para o segundo coeficiente de Pearson, os valores referentes a esses frames serão alocados no vetor de características.

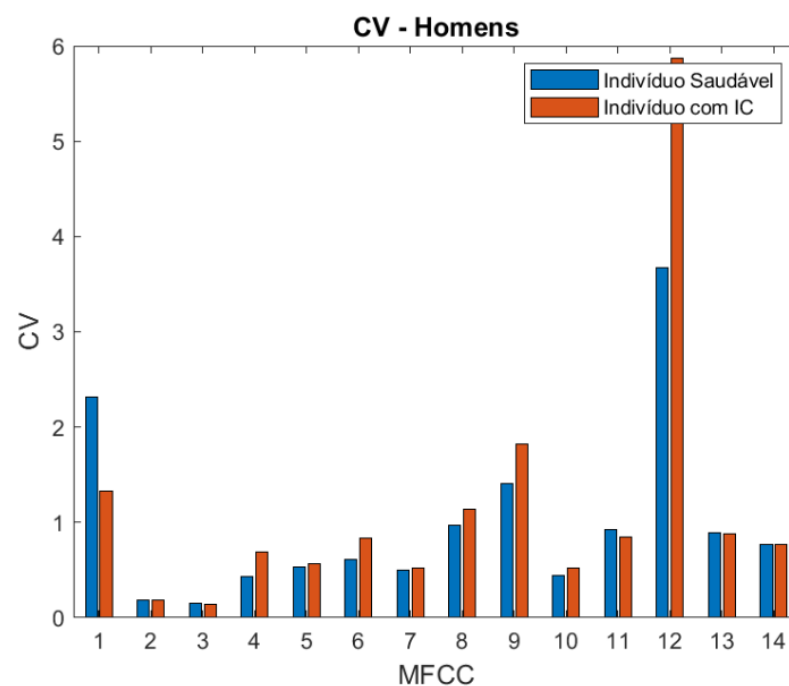


Figura 5.22 - Coeficiente de variação dos MFCCs – Masculino.

A figura 5.22 apresenta uma representação do coeficiente de variação de cada frame dos sinais sonoros das vozes masculinas aqusitadas. Nessa análise, percebeu-se que o coeficiente de variação, apesar de os valores de energia logarítmica (barra 1) e o frame 11 (barra 12) mostrarem valores discrepantes no sinal representado na figura 5.22, esse parâmetro não apresenta um padrão identificável para a caracterização da insuficiência cardíaca ao se analisar os demais sinais.

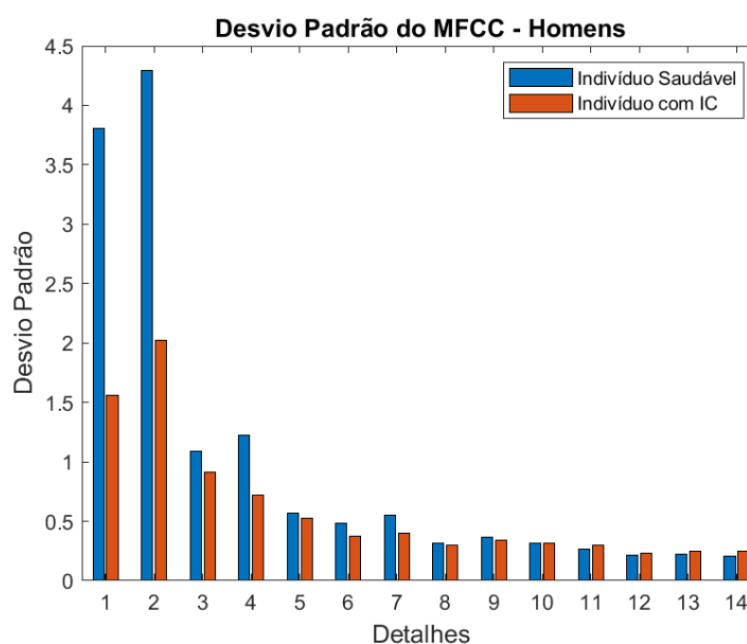


Figura 5.23 - Desvio padrão dos MFCCs – Masculino.

Já para a análise do desvio padrão (figura 5.23), o valor da energia (barra 1) e dos frames 1 e 3 (barras 2 e 4) apresentam valores relevantes para a identificação da IC. Dessa forma, tais parâmetros serão também utilizados para o desenvolvimento da RNA.

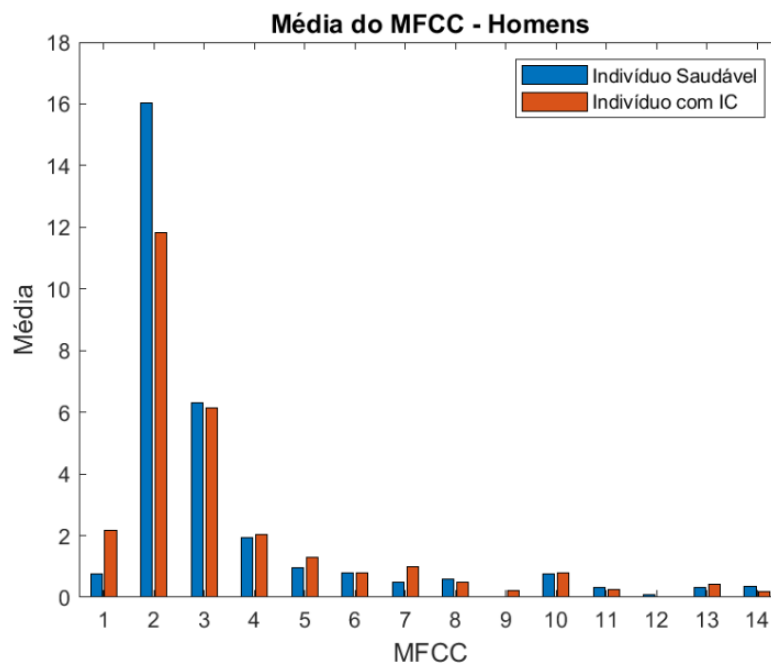


Figura 5.24 - Média dos MFCCs – Masculino.

Para a média de cada frame (figura 5.24), apenas o primeiro frame (barra 2) apresentou importância para o estudo das alterações na voz devido à insuficiência cardíaca. Dessa forma, os demais frames são descartados.

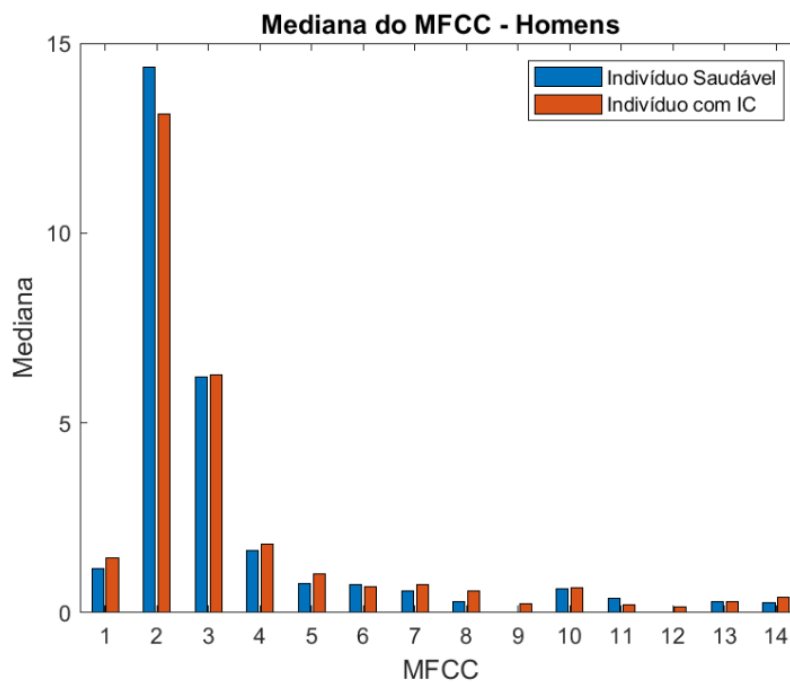


Figura 5.25 - Mediana dos MFCCs – Masculino.

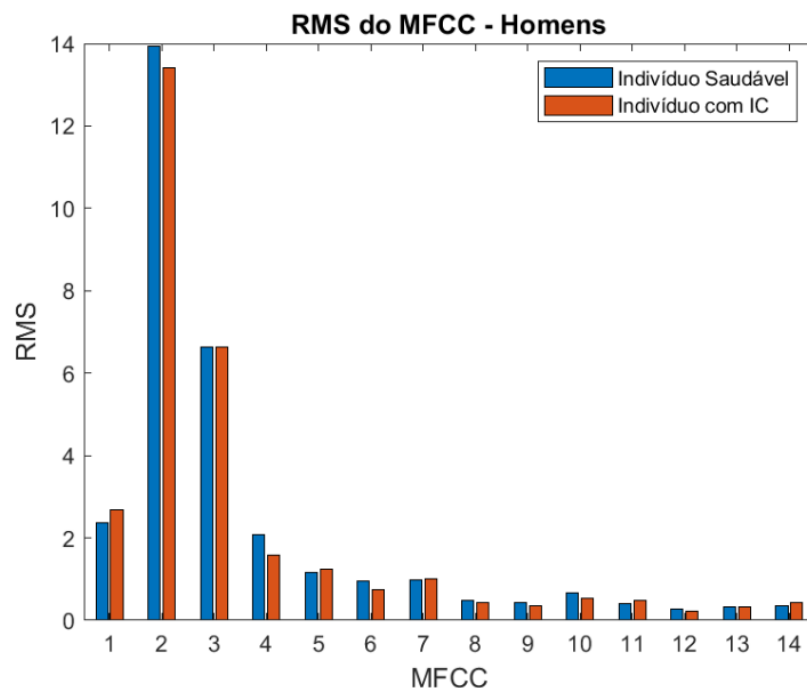


Figura 5.26 - RMS dos MFCCs – Masculino.

As figuras 5.26 e 5.27 apresentam os valores de mediana e RMS, respectivamente. Para esses parâmetros, os valores não são distinguíveis para a diferenciação de um indivíduo com ou sem IC. Dessa forma, nenhum valor do frame para esses dois parâmetros estatístico será selecionado para o vetor de característica.

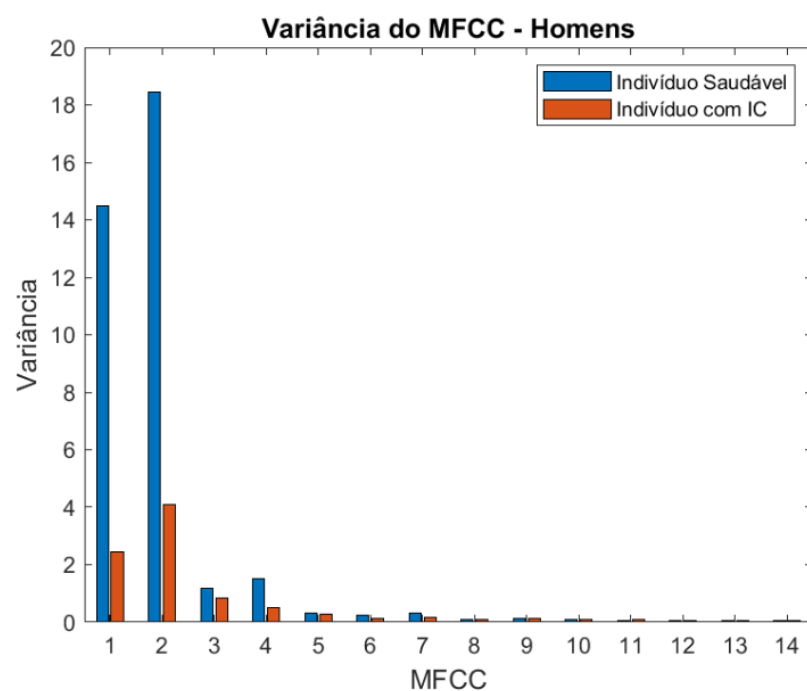


Figura 5. 27 - Variância dos MFCCs – Masculino.

Por fim, a variância dos coeficientes cepstrais para as vozes masculinas (figura 5.27) apresenta relevância nos frames 1 e 3 (barras 2 e 4). Ademais, a energia logarítmica também é um parâmetro para a determinação da IC em indivíduos do sexo masculino. Nesse sentido, a od valores de variância dos homens saudáveis são maiores que os valores para os com IC.

Portanto, a tabela 5.3 apresenta os dados mais pertinentes para identificação de IC de forma mais resumida. Na tabela, esses parâmetros estão marcados com X.

Frame	Energia Log.	1	2	3	4	5	6	7	8	9	10	11	12	13
Amplitude	X	X	-	X	-	-	-	-	-	-	-	-	-	-
Coeficiente de Pearson	-	X	X	X	-	X	-	-	-	-	-	-	-	-
Coeficiente de Variação	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Desvio Padrão	X	X	-	X	-	-	-	-	-	-	-	-	-	-
Média	-	X	-	-	-	-	-	-	-	-	-	-	-	-
Mediana	-	-	-	-	-	-	-	-	-	-	-	-	-	-
RMS	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Variância	X	X	-	X	-	-	-	-	-	-	-	-	-	-

Tabela 5. 3 - Análise dos coeficientes cepstrais – Masculino.

Assim, por meio da tabela 5.3 observa-se que a técnica de análise cepstral dos sinais fornece 14 características diferenciáveis entre indivíduos do sexo masculino saudáveis e com IC.

Para a construção do vetor de características, todos os parâmetros selecionados por meio das técnicas de análise sinais (estatística, FFT e MFCC) são utilizados. Nesse sentido, o vetor de características que irá alimentar a RNA conta com 29 elementos (parâmetros).

5.2 REDE NEURAL PARA OS INDIVÍDUOS DO SEXO MASCULINO

Após a seleção de todas as características relevantes para a determinação do padrão de vozes masculinas saudáveis e doentes, a etapa de desenvolvimento da rede foi iniciada.

Como foram selecionadas 29 características, a camada de entrada da rede terá 29 neurônios de entrada. A camada de saída da RNA terá a quantidade de classes a serem determinadas; no caso do presente estudo, essa camada apresentará 2 neurônios, pois apenas duas classes serão

classificadas: indivíduo doente ou com IC. Por fim, a quantidade de neurônios na camada oculta foi estabelecida por meio de tentativa e erro, de modo a alcançar a maior eficiência possível, visto que não há um padrão para a determinação da quantidade de neurônios para essa camada. Partindo desse princípio, foi aplicada a variação da quantidade de neurônios começando com 5 até chegar em 50, passando de 5 em 5 neurônios. Assim, a tabela 5.4 apresenta os resultados obtidos por meio da aplicação dessa técnica.

Número de Neurônios	Treinamento (%)	Validação (%)	Teste (%)	Eficiência Global (%)
5	80	80	100	83,3
10	90	60	100	86,7
15	95	100	80	93,3
20	95	100	100	96,7
25	80	100	100	86,7
30	100	100	80	96,7
35	100	80	80	93,3
40	80	100	100	86,7
45	95	100	100	96,7
50	100	80	100	86,7

Tabela 5. 4 - Performance da RNA para diferentes quantidades de neurônios na camada oculta – Masculino.

Portanto, a tabela 5.4 mostra que o melhor desempenho obtido pela metodologia proposta neste trabalho foi de 96,7% utilizando 20 neurônios na camada oculta. Além disso, nota-se que os valores correspondentes a 30 e 45 neurônios na camada oculta também apresentam 96,7% de eficiência. Porém, a camada com a menor quantidade de neurônio foi escolhida devido ao seu menor custo computacional. Por fim, a figura 5.28 mostra uma representação da rede desenvolvida para a análise das vozes masculinas.

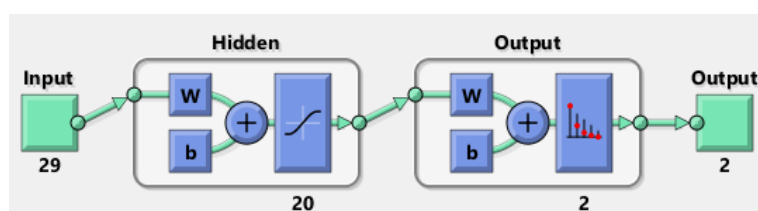


Figura 5.28 - Rede Neural Artificial Desenvolvida – Masculino.

Assim, espera-se que a RNA, representada na figura 5.8, opere com um vetor de características de 29 parâmetros e transfira esses valores por meio da função transferência linear para a próxima camada. Nessa etapa, os dados vão se associar aos pesos e ao bias que, ao passarem pelo operador de soma, adquirem um fator de não linearidade devido à função de transferência tangente hiperbólica. Depois disso, os dados passam pela camada de saída e são associados mais uma vez aos pesos e ao bias por meio do somatório. Por fim, a resposta da rede é gerada pela função de transferência softmax, a qual gera valores estatísticos para cada classe. Dessa forma, a classe que apresenta maior valor estatístico é a classificação gerada pela rede.

5.3 EXTRAÇÃO E SELEÇÃO DAS CARACTERÍSTICAS DAS VOZES FEMININAS

A metodologia empregada para a extração e seleção das características das vozes femininas seguiu o mesmo roteiro aplicado para as vozes masculinas.

5.3.1 Análise Multirresolução dos Sinais

Nessa análise, os sinais provenientes das vozes femininas foram segmentados em 11 faixas de frequências. Essa quantidade de detalhes foi escolhida por permitir uma varredura do sinal em praticamente todas as faixas frequências. Porém, como nem todos os detalhes são importantes para o estudo dos sinais, propõe-se, primeiramente, a análise energética dos dados referentes às vozes femininas. Nesse sentido, a figura 5.29 apresenta a evolução do nível de energia em cada detalhe.

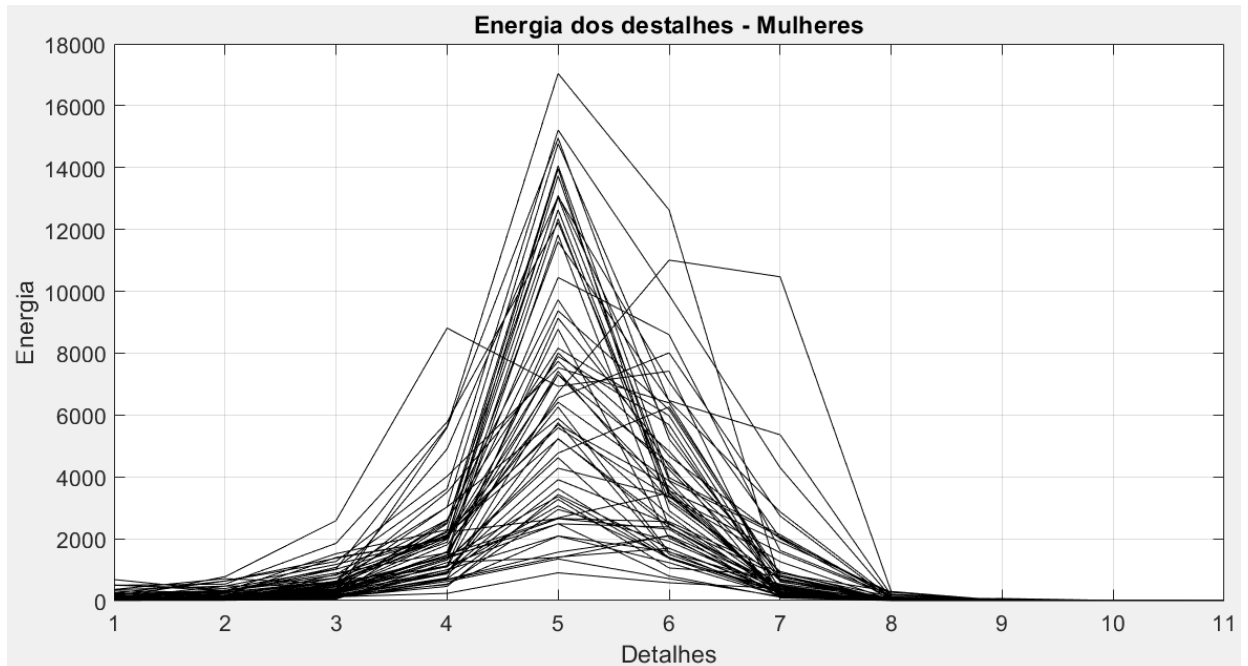


Figura 5.29 - Evolução energética dos detalhes – Feminino.

Com o intuito de diminuir o custo computacional da análise dos sinais e aumentar a eficiência da RNA elaborada, seleciona-se os detalhes com os maiores índices energéticos, como mostrado na figura 5.30.

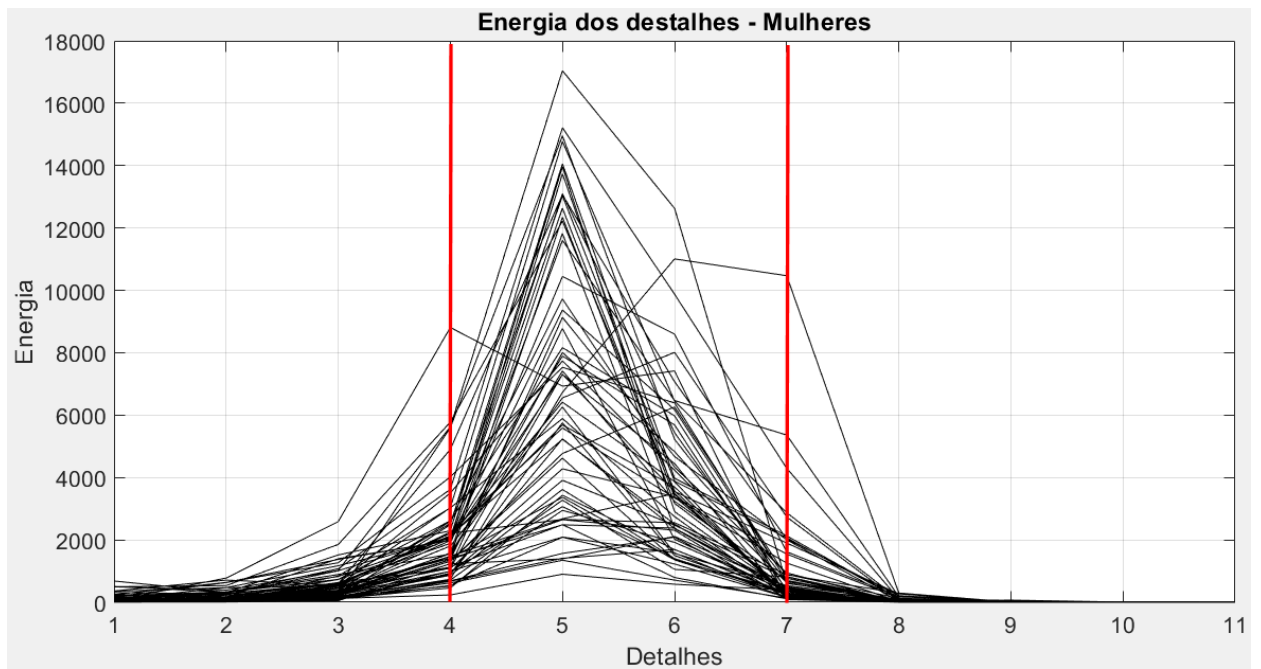


Figura 5.30 - Seleção dos detalhes – Feminino.

Logo, a figura 5.30 mostra que os sinais sonoros das vozes femininas apresentam maior energia nos detalhes 4 (1378,13 – 2756,25 Hz), 5 (689,06 – 1378,13 Hz), 6 (344,53 – 689,06 Hz) e 7 (172,27 – 344,53 Hz). Essa determinação dos detalhes está em consonância com a teoria que aborda o tema, na qual afirma que as cordas vocais femininas vibram em torno de 250 vezes por segundo, isto é, 250 Hz,

Partindo da análise energética, as figuras 5.31 a 5.34 mostram o sinal de uma voz feminina segmentada nos detalhes 4, 5, 6 e 7, respectivamente.

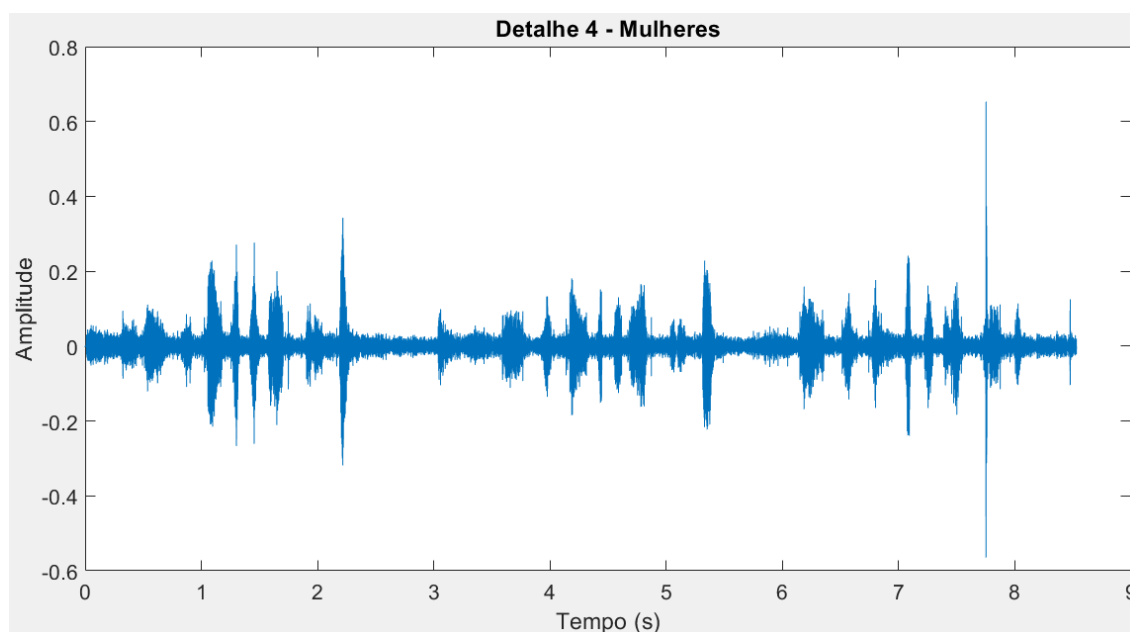


Figura 5.31 - Detalhe 4 – Feminino.

O detalhe 4, mostrado na figura 5.31, representa a variação de frequência que vai de 1378,13 a 2756,25 Hz. Assim, a análise do sinal nessas circunstâncias possibilita o estudo minucioso dessa faixa.

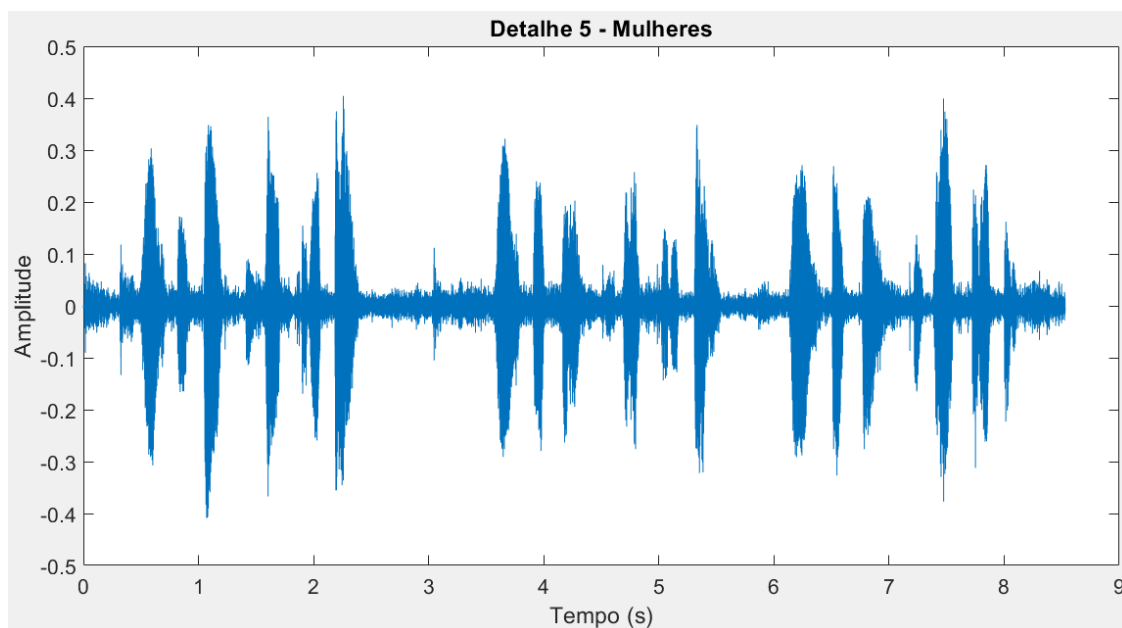


Figura 5.32 - Detalhe 5 – Feminino.

A figura 5.32 representa o detalhe 5 do sinal sonoro de uma voz feminina na faixa 689,06 a 13798,13 Hz. Nesse sentido, as técnicas de análise de sinais estudadas nesta dissertação são aplicadas sobre esse detalhe com o intuito de extrair as características relevantes por meio de estudos estatísticos, do sinal no domínio da frequência e dos coeficientes cepstrais.

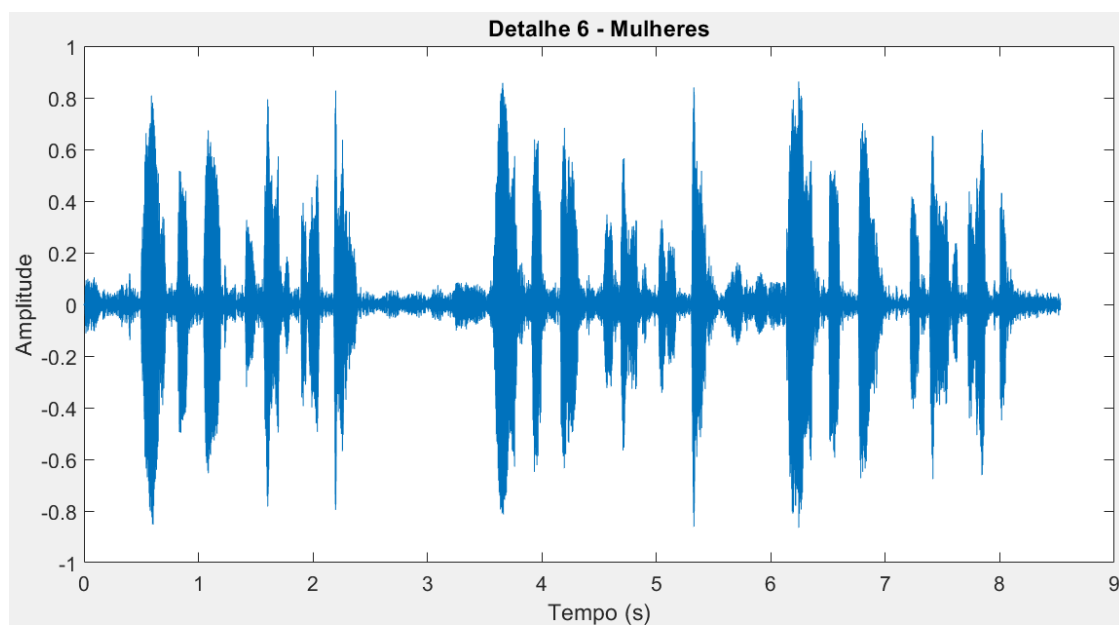


Figura 5.33 - Detalhe 6 – Feminino.

Para o detalhe 6, a figura 5.33 mostra o sinal sonoro da voz feminina de um indivíduo saudável compreendido na faixa de frequência que vai de 344,53 a 689,06 Hz. Assim, é possível extrair e selecionar parâmetros específicos desse detalhe.

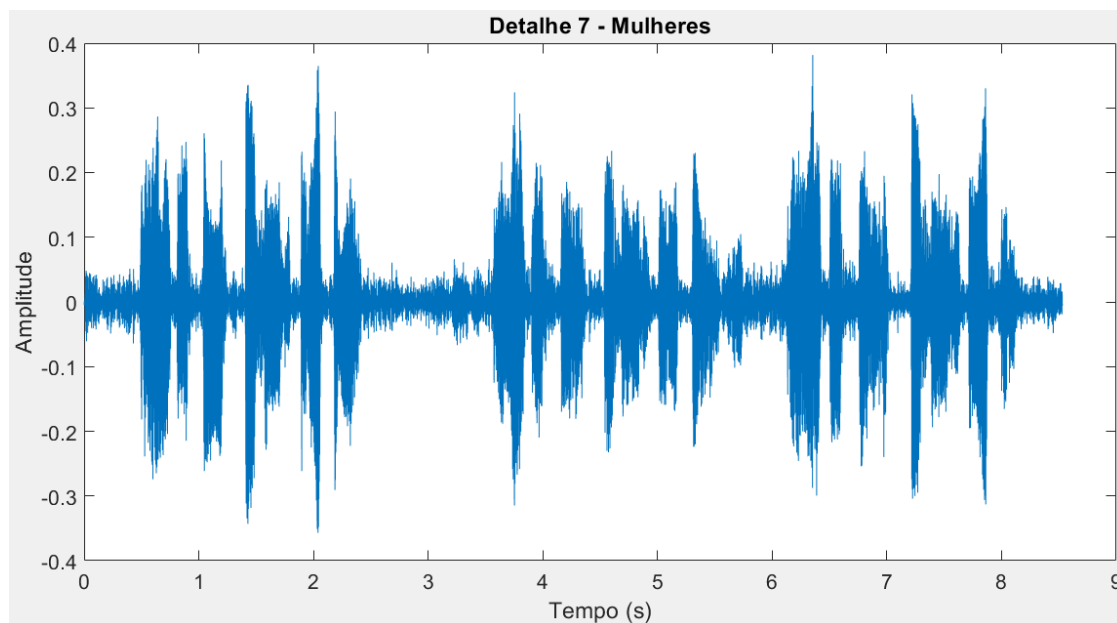


Figura 5.34 - Detalhe 7 – Feminino.

Por último, o detalhe 7, representado pela figura 5.34, apresenta o sinal delimitado entre 172,27 e 344,53 Hz, o que possibilita uma análise mais detalhada dessa faixa de frequência visto que as demais frequências não interferem no estudo.

Após a segmentação dos sinais, inicia-se o processo de análise de cada detalhe por meio das técnicas já apresentadas neste trabalho.

5.3.2 Análise Estatística

Para as vozes femininas, cada um dos detalhes selecionados forneceu os valores de média, mediana, variância, desvio padrão, RMS, amplitude, segundo coeficiente de Pearson e coeficiente de variação, que são exibidos, respectivamente, nas figuras 5.35 a 5.42. As barras dos gráficos apresentam os valores para os homens saudáveis e os com insuficiência cardíaca.

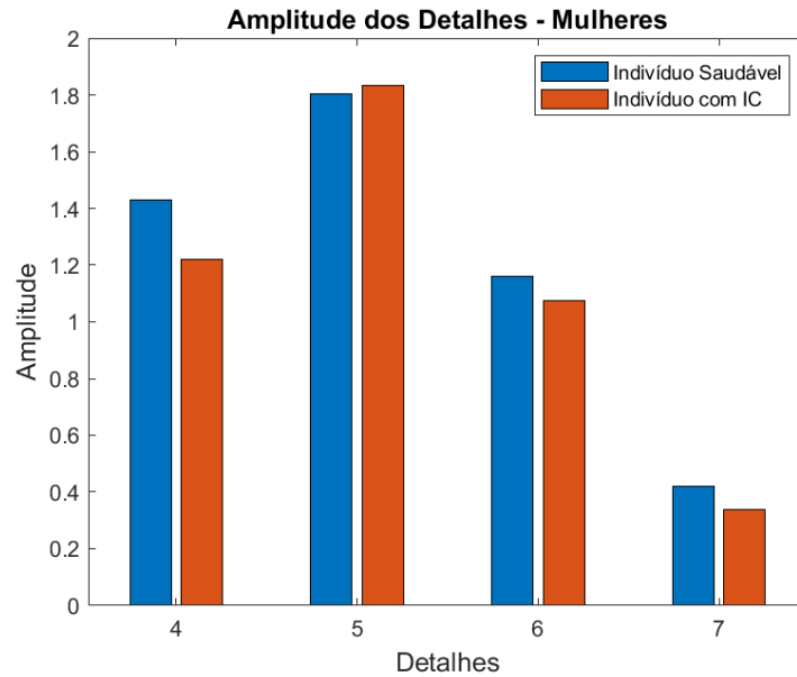


Figura 5.35 - Amplitude dos detalhes – Feminino.

Para a amplitude dos sinais das vozes femininas, a figura 5.35 mostra que em nenhum dos detalhes a amplitude se apresenta como sendo um parâmetro importante para a distinção entre mulheres saudáveis e com IC. Dessa forma, tais valores não serão utilizados como entrada para a elaboração da RNA.

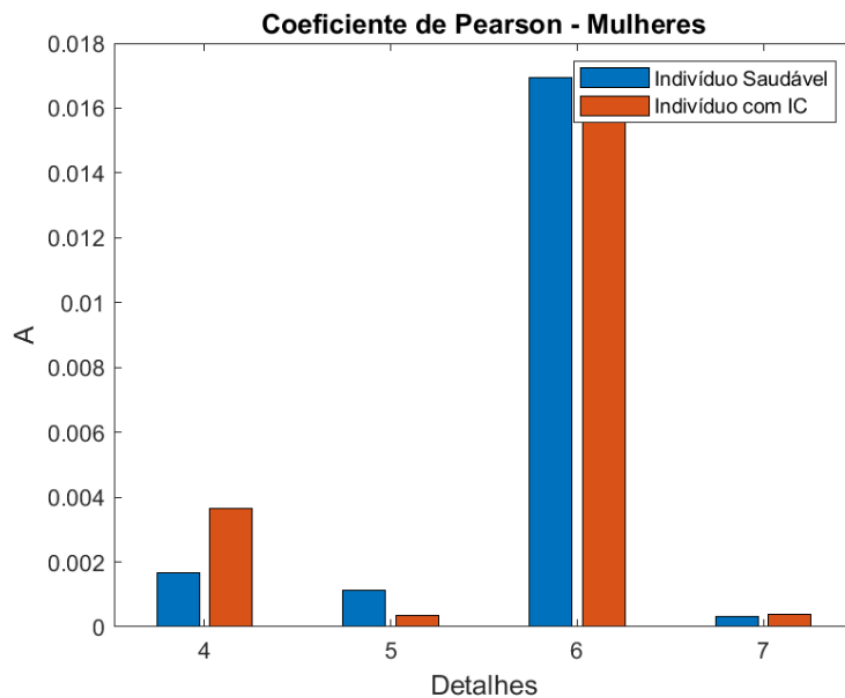


Figura 5. 36 - Coeficiente de Pearson dos detalhes – Feminino.

A figura 5.36 apresenta os valores do segundo coeficiente de Pearson, o qual se caracteriza por ser um valor de dispersão dos sinais. Nesse sentido, verifica-se que os coeficientes de Pearson apresentam uma leve diferenciação no detalhe 4 (344,53 a 689,06 Hz) que se mostra capaz de diferenciar uma mulher saudável de uma com IC.

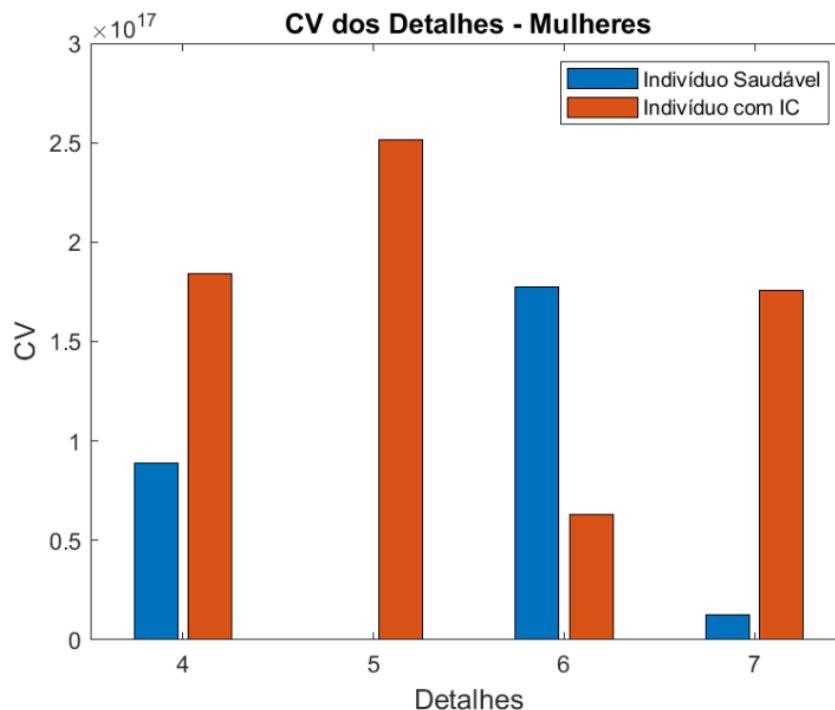


Figura 5. 37 - Coeficiente de variação dos detalhes – Feminino.

Para o coeficiente de variação, embora todos os detalhes apresentem uma grande variação entre os indivíduos saudáveis e com IC, apenas os detalhes 4 e 5 demonstraram seguir um padrão que possibilitou a identificação das mulheres com IC quando uma análise individual de cada sinal é realizada. Nesse sentido, o coeficiente de variação nos detalhes 4 e 5 são, geralmente, superiores para as mulheres doentes.

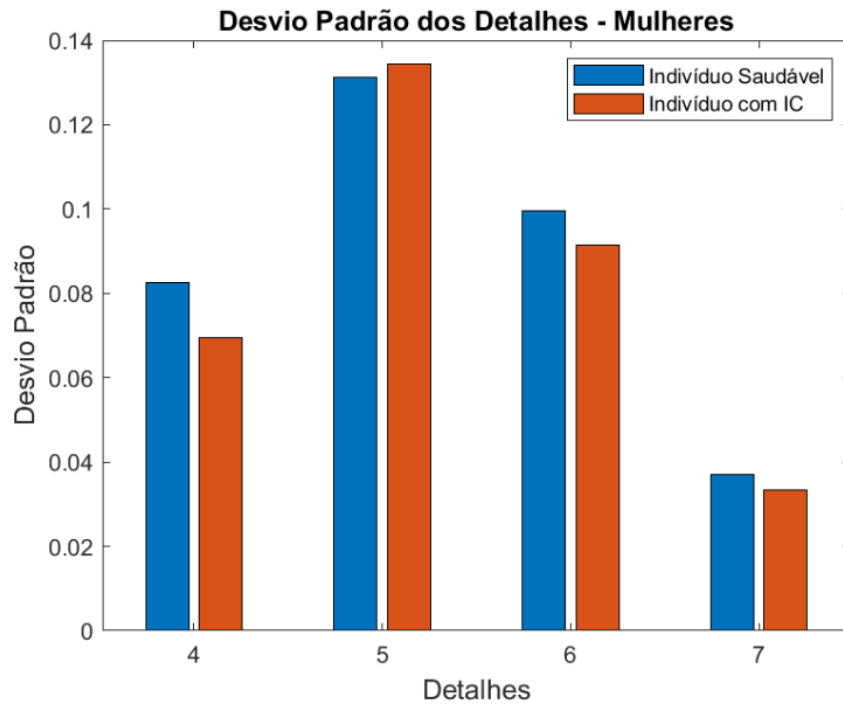


Figura 5.38 - Desvio padrão dos detalhes – Feminino.

Quanto ao desvio padrão dos detalhes das vozes femininas, percebeu-se que não há um padrão identificável entre as mulheres saudáveis e as com IC.

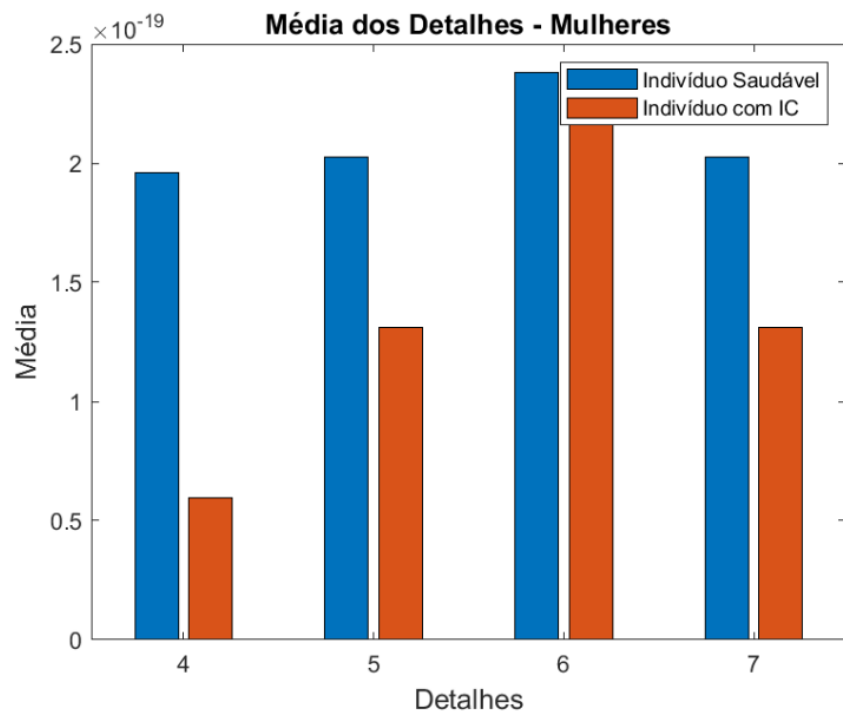


Figura 5.39 - Média dos detalhes – Feminino.

Para a média das faixas em análise, a figura 5.39 mostra que os detalhes 4, 5 e 6 apresentam valores distinguíveis entre os indivíduos saudáveis e doentes. Nesse sentido, em todos esses detalhes, as mulheres saudáveis apresentam valores de média maiores que as mulheres com IC.

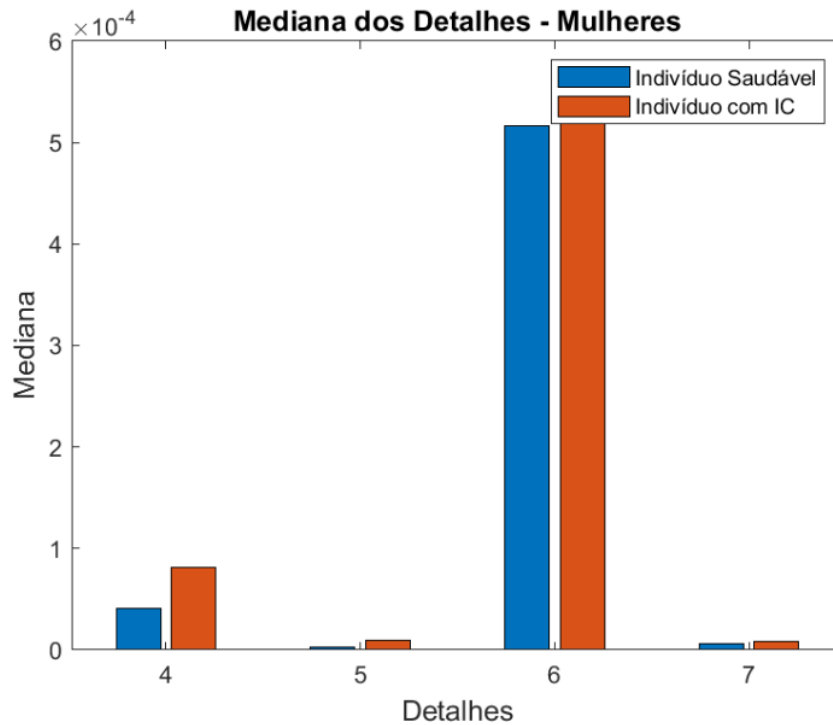


Figura 5. 40 - Mediana dos detalhes – Feminino.

A análise da mediana dos detalhes, figura 5.40, mostra que esse parâmetro apresenta valor significativo no detalhe 6, onde os indivíduos doentes demonstram valores mais elevados

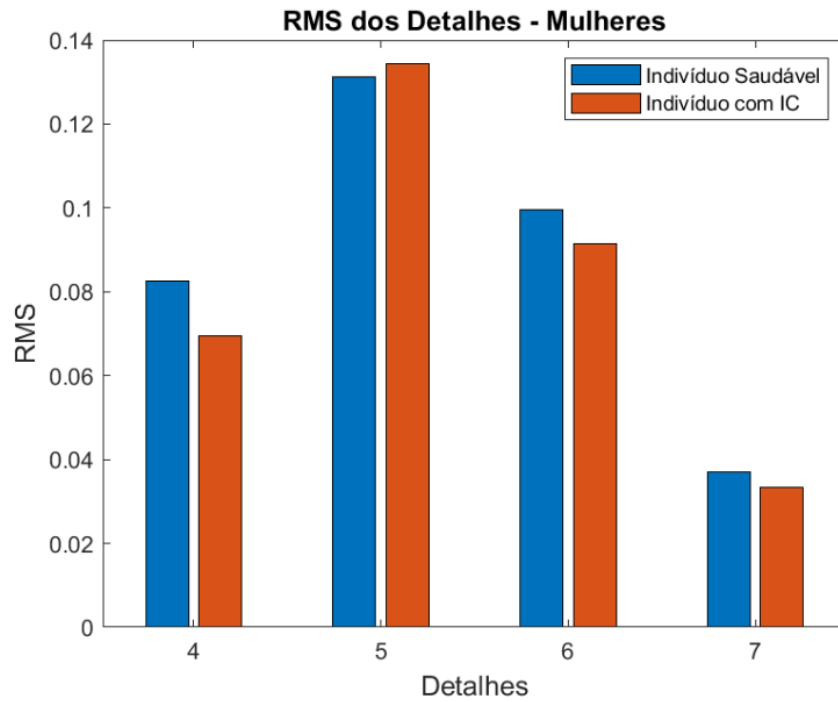


Figura 5. 41 - Valor RMS dos detalhes – Feminino.

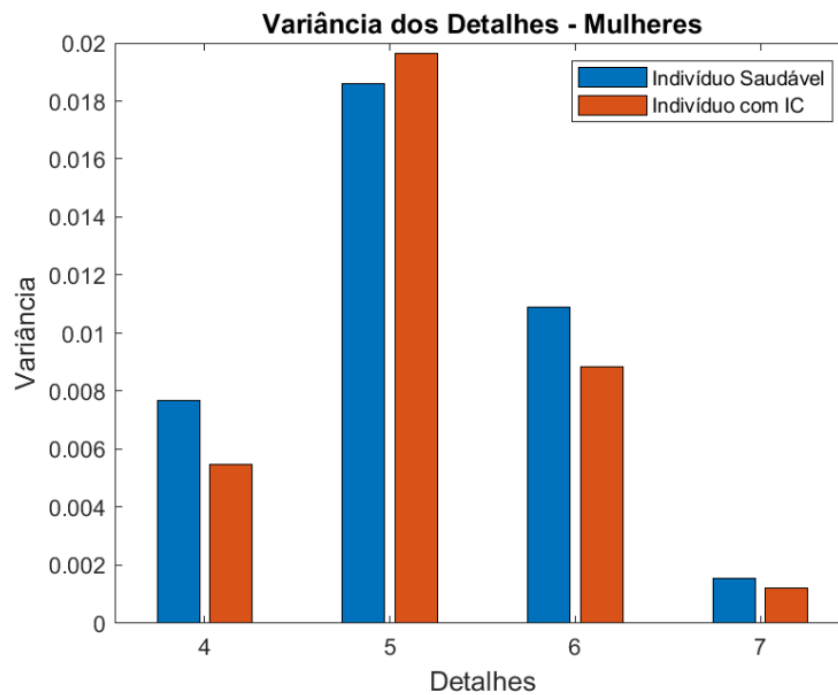


Figura 5. 42 - Variância dos detalhes – Feminino.

As figuras 5.41 e 5.42 mostram os valores RMS e de variância, respectivamente. Assim, os dois parâmetros analisados não apresentaram um padrão identificável para cada classe em estudo.

Por fim, a tabela 5.5 apresenta, de forma sumarizada, as características em função dos detalhes, na qual as propriedades mais relevantes para a diferenciação entre mulheres saudáveis e com IC estão demarcadas com um X.

Detalhe	4	5	6	7
Média	X	X	X	-
Mediana	-	-	X	-
Variância	-	-	-	-
Desvio Padrão	-	-	-	-
RMS	-	-	-	-
Amplitude	-	-	-	-
Coeficiente de Pearson	X	-	-	-
Coeficiente de Variação	X	X	-	-

Tabela 5. 5 - Características relevantes para o desenvolvimento da RNA – Feminino.

Portanto, a tabela 5.5 resume todo o processo de seleção dos parâmetros relevantes para o desenvolvimento do sistema de auxílio ao médico para o diagnóstico da IC em indivíduos do sexo feminino, onde os 7 parâmetros selecionados por meio da análise estatística dos sinais estão marcados com X.

5.3.3 Análise no Domínio da Frequência

Nessa seção, os sinais sonoros das vozes femininas foram analisados no domínio da frequência, onde o estudo foi realizado por meio da aplicação da FFT em cada detalhe selecionado previamente. Dessa forma, a análise dos sinais no domínio da frequência consistiu na seleção do pico predominante e sua respectiva frequência. Assim, as figuras 5.43 a 5.45 mostram um exemplo comparativo dos espectros das vozes femininas saudáveis e doentes.

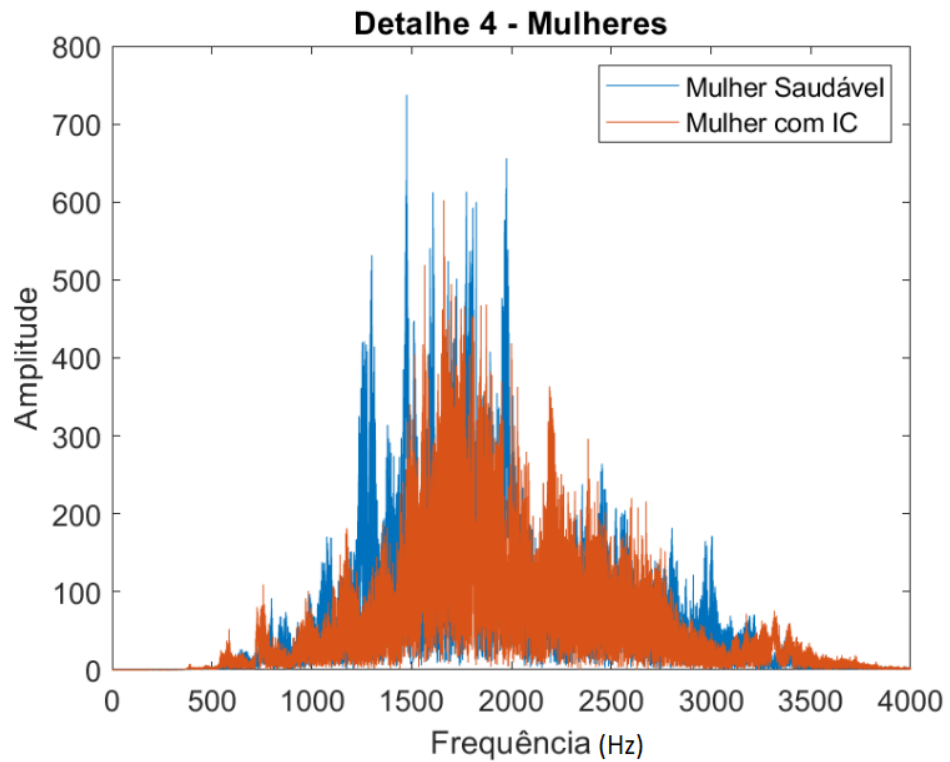


Figura 5. 43 - FFT do detalhe 4 – Feminino.

Para a faixa de frequência que vai de 1378,13 a 2756,25 Hz, representado pelo detalhe 4 e mostrado na 5.43, não foi possível estabelecer um comportamento que possibilite a classificação da insuficiência cardíaca em indivíduos do sexo feminino.

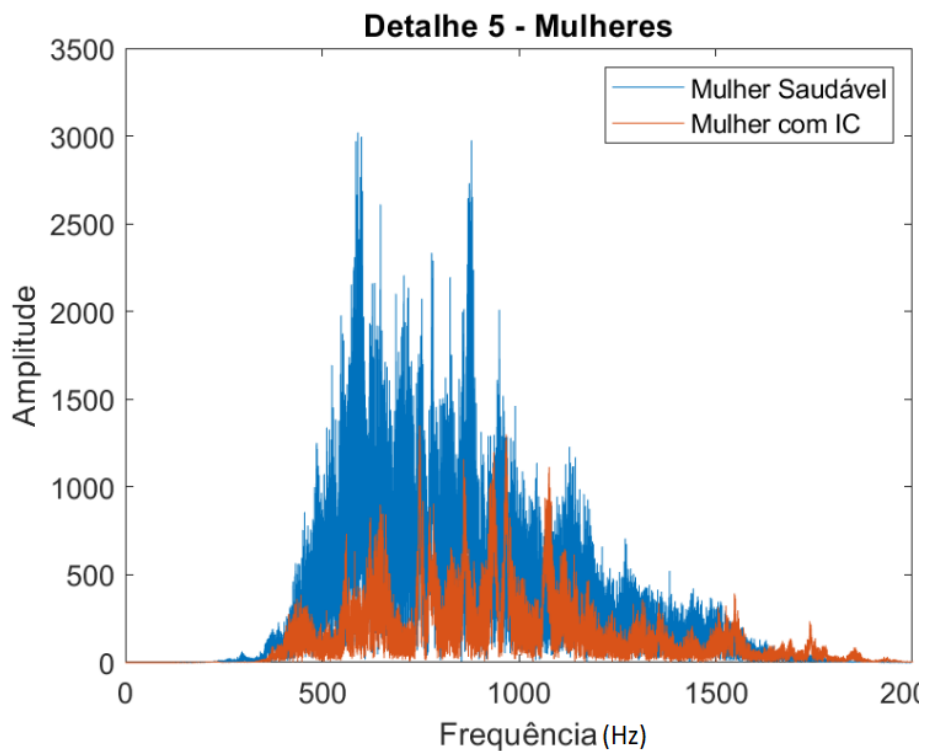


Figura 5. 44 - FFT do detalhe 5 – Feminino.

A figura 5.44 apresenta o padrão do sinal sonoro de uma voz feminina no domínio da frequência e limitado ao detalhe 5. Dessa forma, a assinatura espectral desses sinais está restrita à faixa de frequência que de 689,06 Hz a 1378,13 Hz. Assim, a FFT do detalhe 5 para as mulheres saudáveis apresenta dois picos (900 Hz e 550 Hz) com amplitudes significativamente superiores às amplitudes observadas nas mulheres com IC. Portanto, os picos do detalhe 5 foram utilizados para o desenvolvimento da RNA feminina.

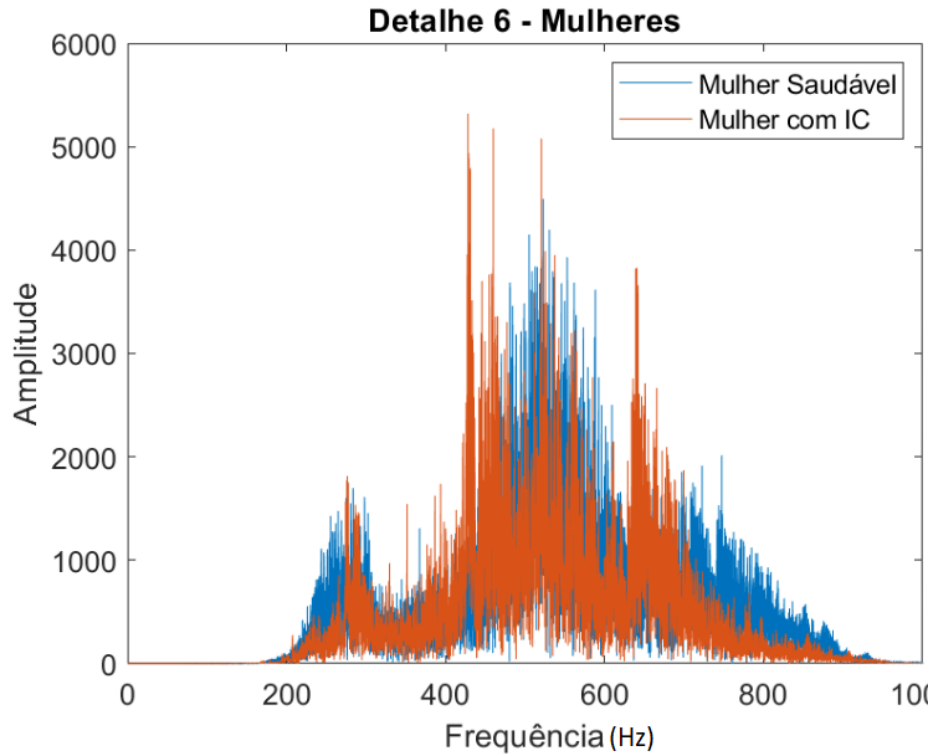


Figura 5.45 - FFT do detalhe 6 – Feminino.

Para a faixa de frequência que vai de 344,53 a 689,06 (detalhe 6), mostrado na figura 5.45, percebe-se que não há pico que se sobressaia perante os demais de forma significativa. Portanto, os picos relativos ao detalhe 6 não serão aplicados para o desenvolvimento da RNA.

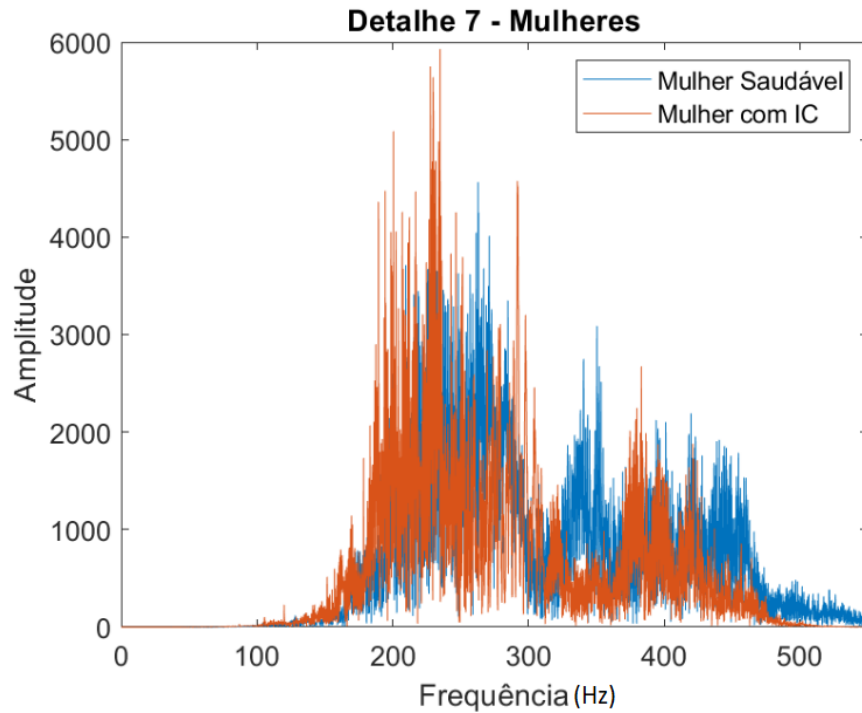


Figura 5.46 - FFT do detalhe 7 – Feminino.

Por fim, o detalhe 7, figura 5.46, apresenta comportamento parecido ao observado no detalhe 6. Assim, a faixa de frequência que corresponde a 172,27 – 344,53 Hz não apresenta diferença entre os indivíduos saudáveis e com IC.

Ademais, a energia de cada faixa de frequência foi analisada e está representada na figura 5.47.

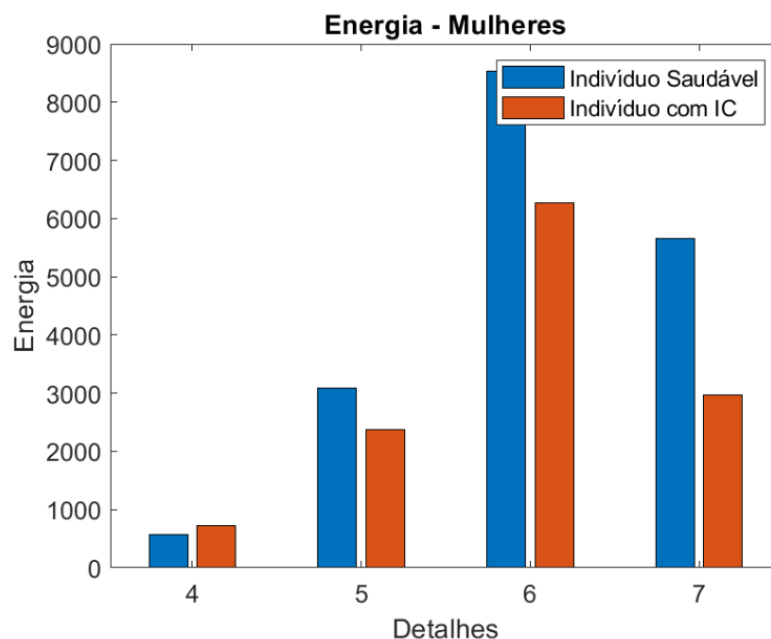


Figura 5. 47 - Energia dos detalhes – Feminino.

A análise energética dos detalhes dos sinais, figura 5.47, não se mostrou importante para esse estudo, visto que não há uma diferença significativa entre os valores encontrados para as mulheres saudáveis e com IC ao se analisar o valor associado a cada sinal individualmente.

A tabela 5.6 mostra de forma mais didática as características selecionadas nessa etapa da pesquisa.

Detalhe	4	5	6	7
Energia	-	-	-	-
Picos (FFT)	-	X	-	-

Tabela 5. 6 - Análise no domínio da frequência – Feminino.

A tabela 5.6 mostra que, por meio da análise dos sinais no domínio da frequência, apenas um parâmetro foi selecionado para compor o vetor de características; tal característica corresponde ao pico proeminente do detalhe 5.

5.3.4 Análise Mel-Cepstral

Em última análise, os valores dos coeficientes cepstrais foram investigados de modo a identificar os valores globais que forneçam características importantes para o desenvolvimento da RNA que identifica o padrão da voz das mulheres saudáveis e das com IC.

Assim como ocorreu com as vozes masculinas, os comprimentos de janela e de sobreposição foram, respectivamente, 1323 e 1103 pontos do sinal. Além disso, a análise cepstral gerou uma matriz de 14 colunas de dados, sendo a primeira referente à energia logarítmica e as demais são relativas aos MFCCs.

Os valores globais examinados nessa etapa são mostrados nas figuras 5.48 a 5.55 e correspondem, respectivamente, aos seguintes parâmetros: amplitude; coeficiente de Pearson; coeficiente de variação (CV); desvio padrão; média; mediana; RMS e variância

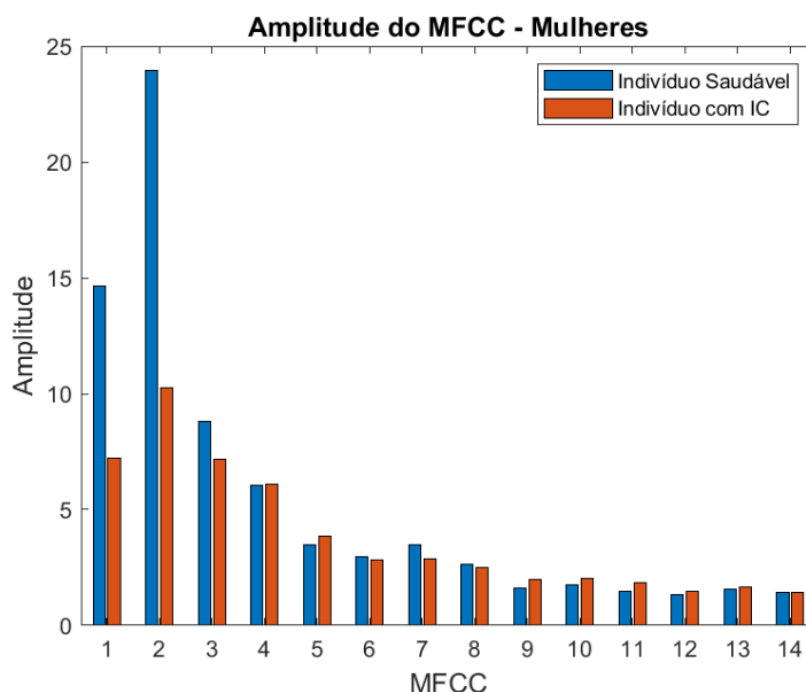


Figura 5.48 - Amplitude dos MFCCs – Feminino.

A figura 5.48 mostra que os valores de amplitude da energia logarítmica e do primeiro frame (barra 2) são distinguíveis ao se comparar as mulheres saudáveis e com IC. Nesse sentido, os valores relativos aos indivíduos saudáveis são substancialmente superiores aos que apresentam IC.

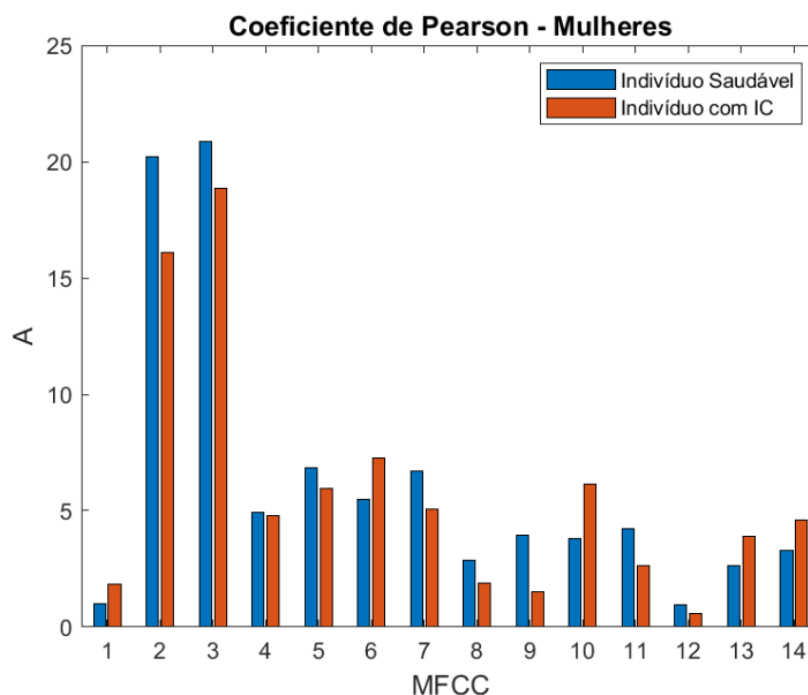


Figura 5. 49 - Segundo coeficiente de Pearson dos MFCCs – Feminino.

O segundo coeficiente de Pearson é relevante nos quadros 1, 2, 6 e 9. Assim, a figura 5.49 mostra que os valores relativos as mulheres saudáveis são superiores para os frames 1 (barra 2), 2 (barra 3) e 6 (barra 7). Por outro lado, os indivíduos que foram diagnosticados com insuficiência cardíaca apresentam um valor mais elevado para o segundo coeficiente de Pearson no quadro 9 (barra 10).

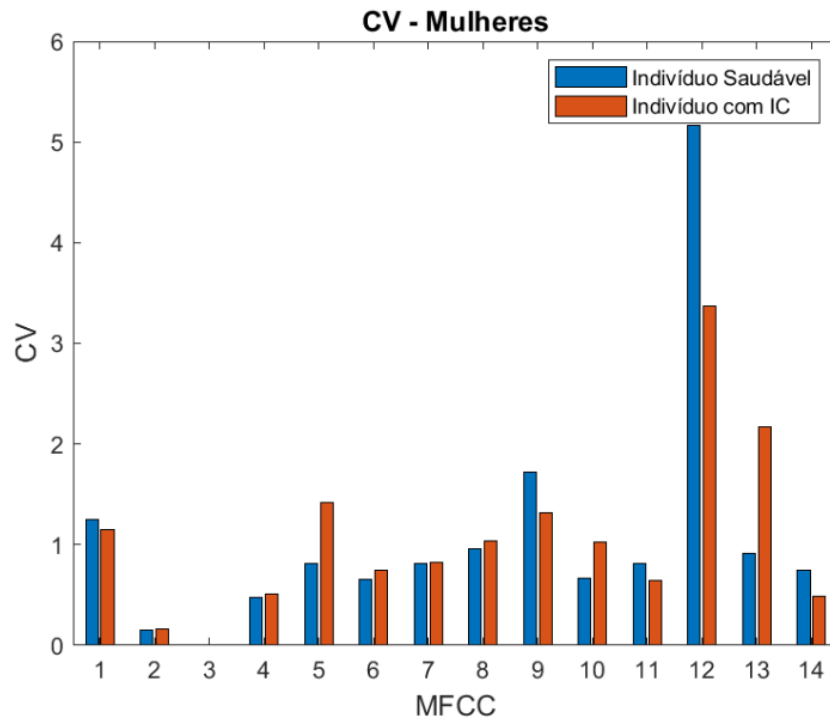


Figura 5. 50 - Coeficiente de Variação dos MFCCs – Feminino.

A figura 5.50 representa o coeficiente de variação de cada frame dos sinais sonoros das vozes femininas. Assim, observa-se que esse parâmetro apresenta uma discrepância entre seus valores no frame 11 (barra 12), porém a análise individual de cada voz não apresentou um padrão consistente a ponto de identificar um padrão para as vozes de indivíduos com IC. Portanto, não se utilizou o coeficiente de variação dos coeficientes cepstrais das vozes femininas para o desenvolvimento da RNA.

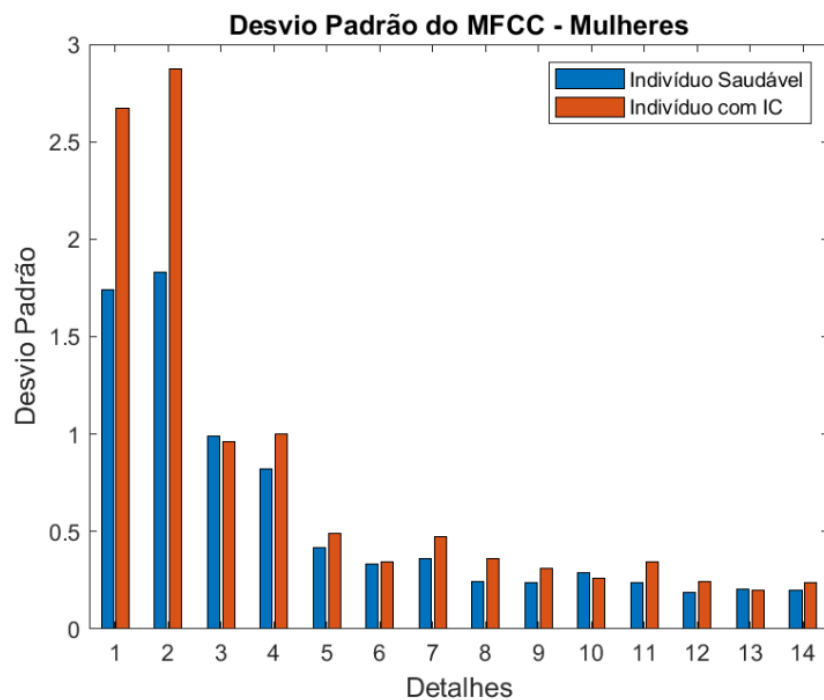


Figura 5.51 - Desvio Padrão dos MFCCs – Feminino.

Para o desvio padrão dos coeficientes cepstrais, vê-se, por meio da figura 5.51, que o primeiro quadro (barra 2) e a energia logarítmica (barra 1) apresentam valores mais elevados para as mulheres com IC.

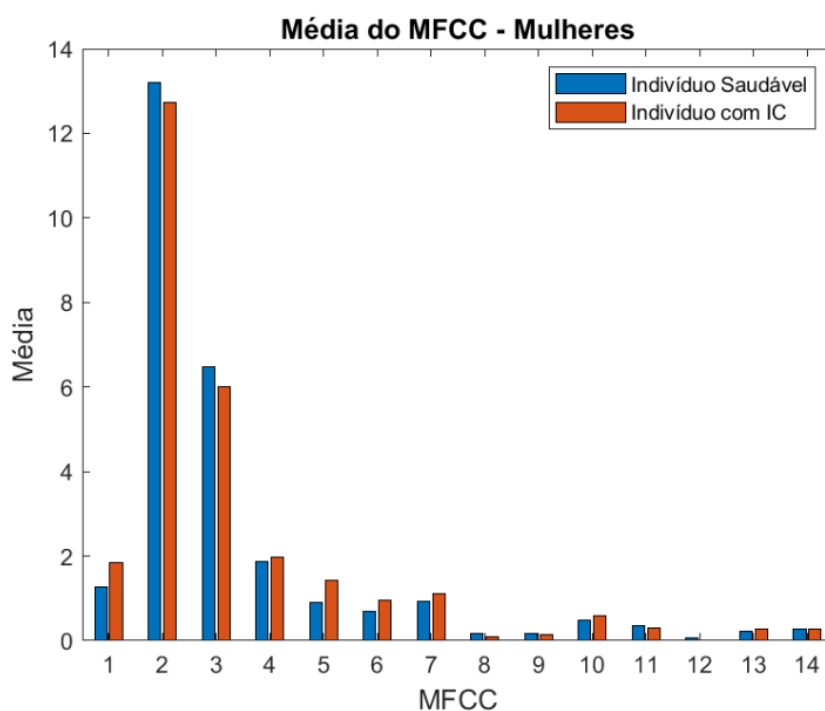


Figura 5.52 - Média dos MFCCs – Feminino.

A figura 5.52 mostra os valores de média, onde seus valores não apresentam relevância para a determinação da IC nas mulheres.

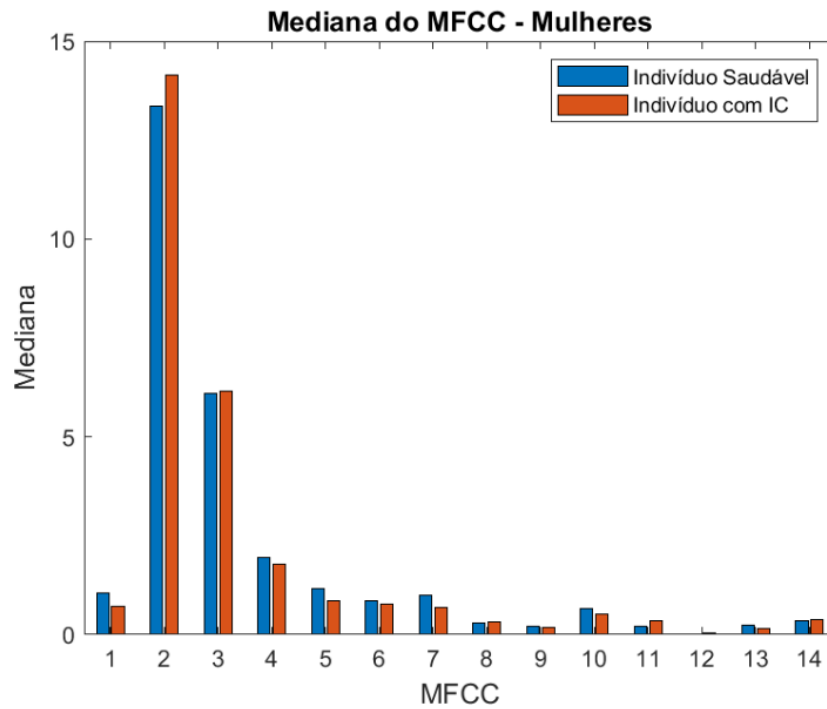


Figura 5.53 - Mediana dos MFCCs – Feminino.

Assim como na média, os valores de mediana também não são utilizados para o desenvolvimento da RNA, como mostra a figura 5.53.

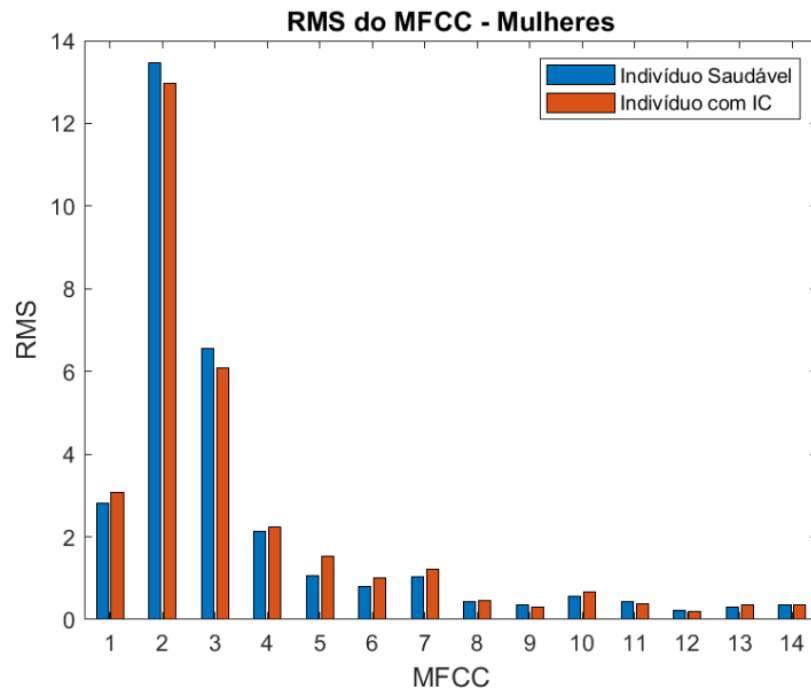


Figura 5.54 - Valor RMS dos MFCCs – Feminino.

Além da média e da mediana, os valores de RMS, mostrado na figura 5.54, também não apresentam um padrão que diferencie as mulheres saudáveis das doentes.

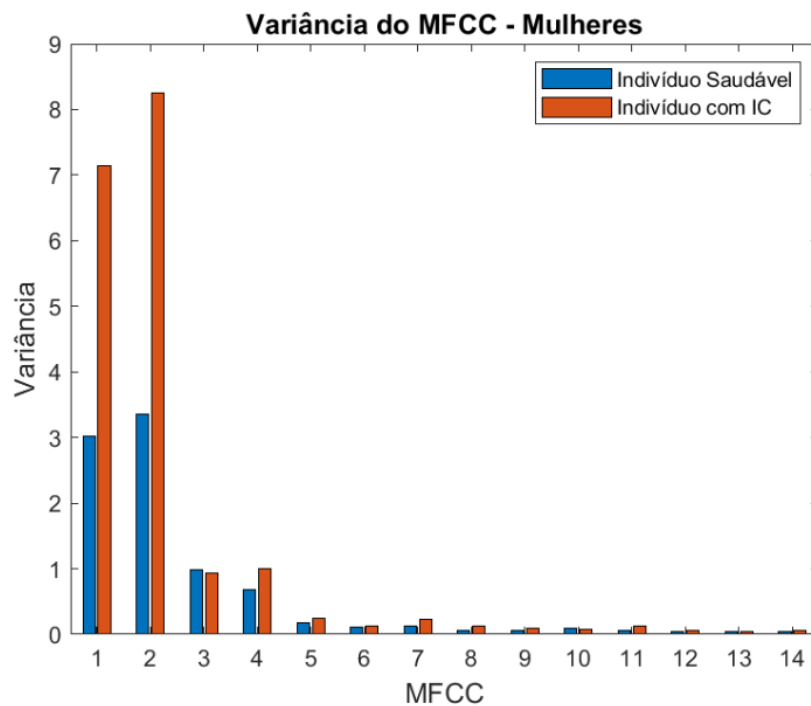


Figura 5.55 - Variância dos MFCCs – Feminino.

Por fim, foi analisado a variância dos MFCCs. Nesse sentido, observa-se, por meio da figura 5.55, que a energia logarítmica (barra 1) e o primeiro frame (barra 2) são significativamente superiores para as mulheres com IC. Assim, esse parâmetro foi aplicado no desenvolvimento da rede neural.

A tabela 5.7 mostra os parâmetros importantes para a criação da rede obtidos por meio da análise cepstral dos sinais.

Frame	Energia Log.	1	2	3	4	5	6	7	8	9	10	11	12	13
Amplitude	X	X	-	-	-	-	-	-	-	-	-	-	-	-
Coeficiente de Pearson	-	X	X	-	-	X	-	-	-	X	-	-	-	-
Coeficiente de Variação	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Desvio Padrão	X	X	-	-	-	-	-	-	-	-	-	-	-	-
Média	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Mediana	-	-	-	-	-	-	-	-	-	-	-	-	-	-
RMS	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Variância	X	X	-	-	-	-	-	-	-	-	-	-	-	-

Tabela 5. 7 - Análise dos coeficientes cepstrais – Masculino.

A tabela 5.7 destaca com um X os parâmetros selecionados para servir de entrada da RNA desenvolvida. Dessa forma, da análise cepstral foi possível a extração e seleção de 10 parâmetros que vão integrar do vetor de características.

5.4 REDE NEURAL PARA OS INDIVÍDUOS DO SEXO FEMININO

Após a seleção dos parâmetros importantes para a determinação do padrão de vozes femininas saudáveis e doentes, deu-se início ao desenvolvimento da RNA voltada à identificação desse padrão de forma automática.

Para as mulheres, foram selecionadas 18 características. Dessa forma, a camada de entrada da RNA terá 18 neurônios. A quantidade de neurônios na camada de saída da rede será determinada pela quantidade de classes a serem classificadas; ou seja, tal camada tem contará com dois neurônios. Por último, a quantidade de neurônios na camada oculta foi estabelecida por meio de tentativa e erro, assim como ocorreu no desenvolvimento da RNA voltada para as vozes

masculinas. Assim, a tabela 5.8 apresenta os resultados obtidos por meio da aplicação dessa técnica.

Número de Neurônios	Treinamento (%)	Validação (%)	Teste (%)	Eficiência Global (%)
5	90	100	80	90
10	95	100	60	90
15	100	100	80	96,7
20	90	80	80	86,7
25	85	80	100	86,7
30	100	80	80	93,3
35	95	80	100	93,3
40	85	60	100	83,3
45	95	80	100	93,3
50	100	100	80	96,7

Tabela 5. 8 - Performance da RNA para diferentes quantidades de neurônios na camada oculta – Feminino.

Assim, a tabela 5.8 mostra que as redes com 15 e 50 neurônios são as que apresentam maior eficiência global, 96,7%. Dessa forma, por questão de custo computacional, escolheu-se a RNA com 15 neurônios na camada oculta. Por fim, a figura 5.56 mostra uma representação da rede criada para a avaliação das vozes femininas.

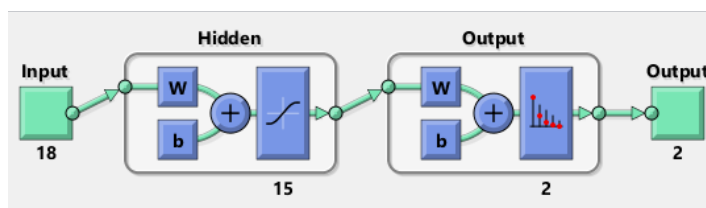


Figura 5. 56 - Rede Neural Artificial Desenvolvida – Feminino.

Da RNA desenvolvida, representada na figura 5.56, espera-se que ela trabalhe com um vetor de características de 18 parâmetros e transfira esses valores por meio da função transferência linear para a próxima camada. Após associar os dados aos pesos e ao bias, a próxima etapa consiste em adicionar a não linearidade a esses dados por meio da função de transferência tangente hiperbólica. Por último, a resposta da rede é gerada pela camada de saída, a qual opera utilizando a função de transferência softmax e gera valores estatísticos para cada classe. Dessa forma, a classe que apresenta maior valor estatístico é a classificação gerada pela rede.

5.5 DESENVOLVIMENTO E FUNCIONAMENTO DO APLICATIVO

Como a utilização do programa em sua forma usual de código não é amigável para a sua aplicação diária pelos profissionais de saúde, optou-se por desenvolver uma interface gráfica para facilitar a interação do programa com o usuário.

O software desenvolvido funciona com um aplicativo comum, no qual um ícone na área de trabalho do computador fornece acesso à sua interface. O aplicativo foi gerado por meio do Matlab utilizando da função ‘*deploytool*’ que funciona como um compilador de aplicativos. A figura 5.57 exibe a interface mostrada pelo aplicado ao ser acessado.

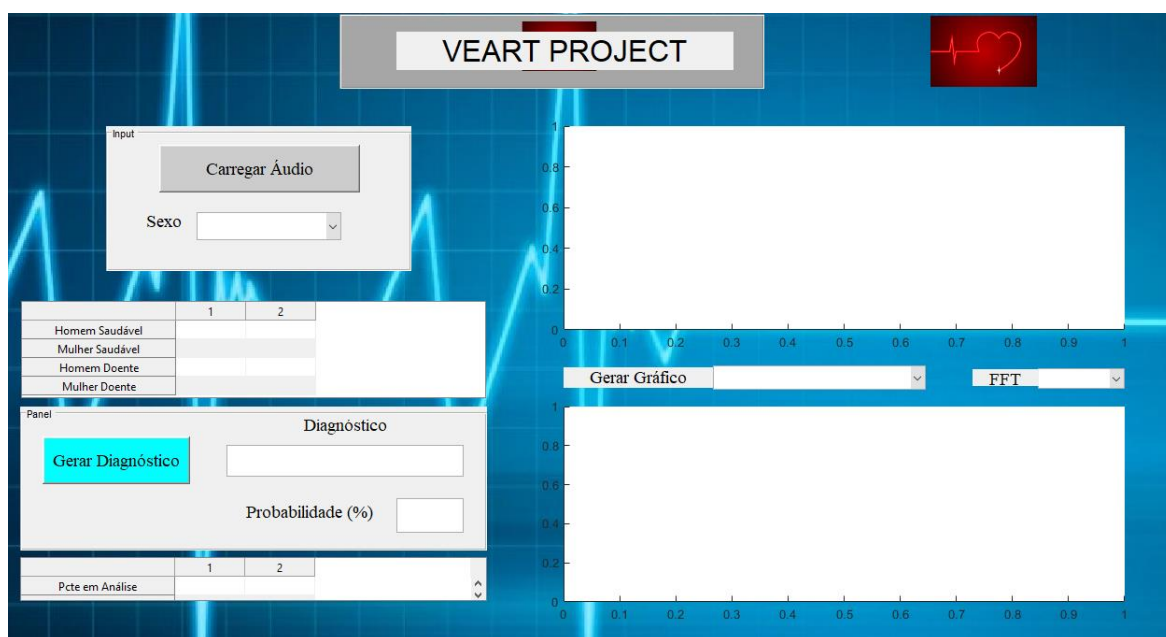


Figura 5. 57 - Aplicativo Desenvolvido para Análise das Vozes – Feminino.

O software desenvolvido apresenta diversas funcionalidades que podem ser acessadas em sua tela inicial. Na parte superior esquerda da interface, foi inserido um botão de carregar áudio, no qual o usuário pode selecionar um áudio que esteja em qualquer diretório do computador. Logo abaixo desse botão há um campo destinado à seleção do sexo do indivíduo em análise, podendo ser masculino ou feminino.

Após a etapa de carregar o áudio e selecionar o sexo da pessoa em estudo, o aplicativo vai gerar automaticamente os valores padrão de média, mediana, desvio padrão e amplitude das

peessoas saudáveis e com insuficiência cardíaca, de ambos os sexos, a fim de gerar dados de comparação. Ao mesmo tempo, o gráfico do sinal sonoro da voz no domínio do tempo será exibido no painel fixado na parte superior direita da interface.

Por fim, no painel localizado na parte inferior esquerda do aplicativo se encontra o botão de ‘Gerar Diagnóstico’ no qual, ao ser pressionado, fornece a classificação do indivíduo em ‘Paciente com IC’ ou ‘Indivíduo Saudável’. Ademais, a probabilidade de acerto é exibida logo abaixo do diagnóstico. Além disso, para efeitos de comparação, uma tabela com valores de média, mediana, desvio padrão e amplitude são mostrados em uma tabela no canto inferior esquerdo do aplicativo. Por último, há a opção de gerar gráficos na parte inferior direita da interface. Nesse sentido, as opções de gráficos são: média, mediana, desvio padrão, amplitude e transformada de Fourier (detalhes 5 e 7). Assim, a figura 5.58 apresenta uma visualização da classificação gerada pelo aplicativo de um indivíduo masculino com insuficiência cardíaca previamente diagnosticada.

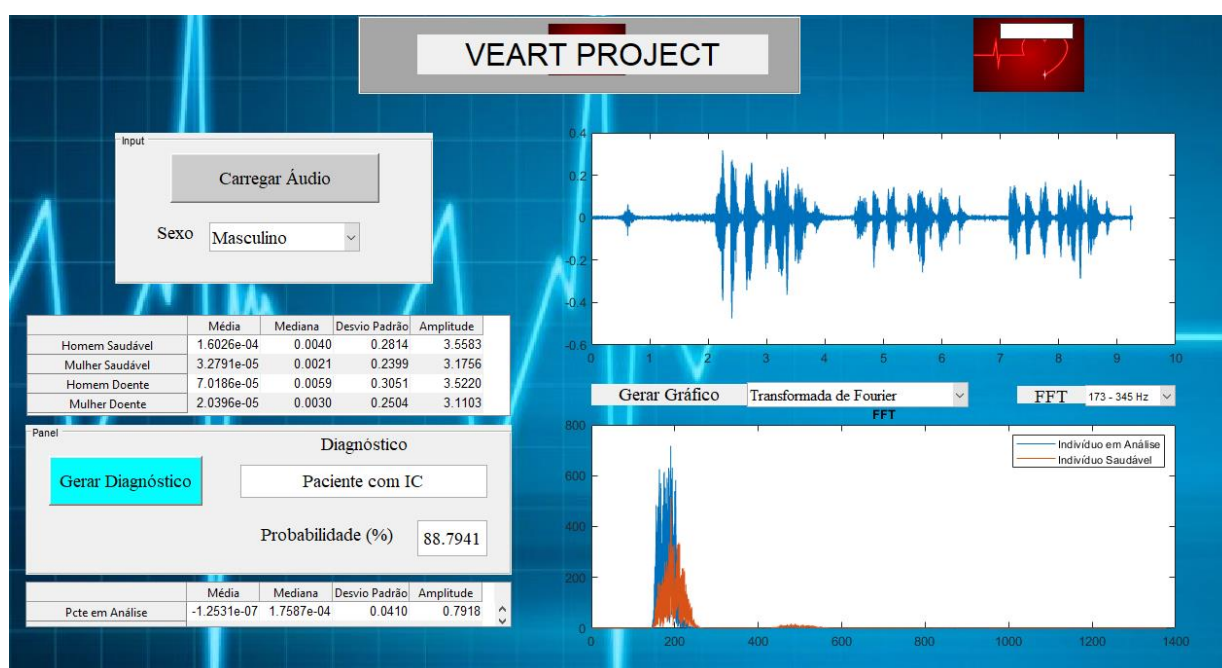


Figura 5. 58 - Classificação Gerada pelo Aplicativo – Feminino.

5.6 ACURÁCIA, SENSIBILIDADE E ESPECIFICIDADE DO SISTEMA

A fim de testar a possível aplicabilidade e confiabilidade do aplicativo desenvolvido, alguns testes extras foram realizados. Partindo dessa perspectiva, extraiu-se os valores de acurácia, sensibilidade e especificidade por meio das equações 5.1, 5.2 e 5.3, respectivamente.

$$Ac = \frac{a+d}{n} \quad (5.1)$$

$$S = \frac{a}{a+c} \quad (5.2)$$

$$E = \frac{d}{b+d} \quad (5.3)$$

Nessas equações, a representa o verdadeiro positivo, ou seja, um indivíduo com insuficiência foi corretamente diagnosticado com a doença; b é o falso positivo – uma pessoa saudável foi, erroneamente, classificada como doente; c representa o falso negativo, onde uma pessoa doente foi identificada como sendo saudável; e d é o verdadeiro negativo – um indivíduo foi, corretamente, classificado com sendo saudável. Por último, n é o total de testes realizados.

Assim, ao analisar as equações 5.1 a 5.3, percebe-se que a acurácia se refere à probabilidade de o teste fornecer resultados corretos, sejam os indivíduos saudáveis ou não. Por outro lado, a sensibilidade representa a probabilidade do verdadeiro positivo ser mostrado pelo aplicativo. Finalmente, a especificidade fornece a probabilidade de ocorrer o verdadeiro negativo.

Para a determinação da acurácia, da sensibilidade e da especificidade do software elaborado, utilizou-se as vozes de 14 homens saudáveis e 40 homens com insuficiência cardíaca; para as mulheres, foram aplicadas as vozes de 22 saudáveis e 13 doentes. Os resultados obtidos estão expostos na tabela 5.9.

Sexo / Resultados	Verdadeiro Positivo (a)	Falso Positivo (b)	Falso Negativo (c)	Verdadeiro Negativo (d)
Homens	38	1	2	13
Mulheres	10	2	3	20
Total	48	3	5	33

Tabela 5. 9 - Resultados Obtidos dos Testes – Feminino.

Assim, ao utilizar os valores mostrados na tabela 5.9 nas equações 5.1 a 5.3, chega-se aos valores de acurácia, sensibilidade e especificidade para cada sexo, conforme exposto na tabela 5.10.

Sexo / Resultados	Acurácia (%)	Sensibilidade (%)	Especificidade (%)
Homens	94,34	95	93
Mulheres	87,88	77	91
Total	91,86	88,1	92,1

Tabela 5. 10 - Acurácia, Sensibilidade e Especificidade do sistema – Feminino.

Ao analisar os resultados relativos aos indivíduos do sexo masculino, percebe-se que estes mostraram-se melhores quando comparados aos resultados observados nos indivíduos do sexo feminino. Assim, tal diferença observada nos resultados pode ser explicada pela qualidade da análise em cada grupo dados, visto que, na etapa de estudo das características dos indivíduos, foram analisadas 84 vozes masculinas e 57 vozes femininas.

Os resultados mostrados na tabela 5.10 apresentam-se, a princípio, animadores. Porém, tais valores precisam ser contextualizados para que algumas conclusões possam ser feitas. Assim, com esse objetivo, os valores obtidos pelo aplicativo desenvolvido foram comparados com algumas técnicas já consagradas para o diagnóstico de IC. Nesse sentido, a tabela 5.11 apresenta valores de acurácia, sensibilidade e especificidade alcançadas pela análise dos Peptídeos Natriuréticos B (BNP e NT-ProBNP). Esses dados foram encontrados no relatório de recomendação de maio de 2018 da Comissão Nacional de Incorporação de Tecnologias do SUS (CONITEC). Dessa forma, para fins de comparação, a primeira linha da tabela 5.11 representa os valores obtidos nessa pesquisa e as demais representam as técnicas apresentadas no relatório.

Número de Estudos	Análise (%)	Acurácia (%)	Sensibilidade (%)	Especificidade (%)
1	Voz / RNA	91,86	88,1	92,1
28	BNP	80	79	72
32	NT-proBNP	80	84	55

Tabela 5. 11 - Análise Comparativa das técnicas empregadas para o diagnóstico de IC – Feminino.

Portanto, percebe-se que os valores de acurácia, sensibilidade e especificidade obtidos pela presente pesquisa são superiores aos alcançados pelos estudos apresentados pela CONITEC.

Porém, é importante enfatizar que a técnica de análise de voz para a identificação da insuficiência cardíaca ainda precisa de um volume maior de testes, em diferentes cenários, de modo a comprovar sua eficácia e eficiência. Apesar dessa ressalva, o estudo se mostra bastante promissor.

5.7 CONCLUSÃO

O capítulo 5 apresenta os resultados obtidos nesse trabalho, por meio das técnicas e metodologia explicada nos capítulos anteriores.

Para os homens, foram selecionadas 29 características que diferenciaram os indivíduos saudáveis dos indivíduos com IC. A RNA dedicada ao sexo masculino apresentou uma eficiência global de 96,7%.

Por outro lado, utilizando as técnicas de análise de sinais, foi possível selecionar 18 características que diferenciaram as mulheres saudáveis das mulheres com IC. A RNA dedicada ao sexo feminino também apresentou uma eficiência global de 96,7%.

Por último, o aplicativo de auxílio ao diagnóstico da IC foi desenvolvido por meio do Matlab. Esse software teve como principal objetivo facilitar o uso dos programas computacionais pelos profissionais da saúde. Além disso, o sistema apresentou valores de 91,86%; 88,1% e 92,1% de acurácia, sensibilidade e especificidade, respectivamente. Esses valores mostraram-se promissores quando comparados com outras técnicas (BNP e NT-proBNP) já estabelecidas no mercado.

CAPÍTULO VI

CONCLUSÕES

Nesse capítulo, apresenta-se as conclusões realizadas por meio da análise dos resultados obtidos no capítulo anterior. Ainda, é mostrada algumas sugestões de possíveis trabalhos futuros que possam aprimorar o atual projeto.

6.1 CONCLUSÕES

Com base nos resultados obtidos, conclui-se que:

- a) A análise sonora da voz de indivíduos, utilizando as técnicas de estudo de sinais nos domínios do tempo (estatística), frequência (FFT) e tempo-frequência (DWT), possibilitou a diferenciação entre indivíduos com insuficiência cardíaca de pacientes saudáveis.
- b) Dentre as técnicas utilizadas, a análise mel-cepstral se mostrou uma ferramenta poderosa para o estudo das vozes alteradas por patologias cardiorrespiratórias, devido a grande quantidade de dados que essa técnica fornece à análise dos sinais.
- c) No processo de elaboração da pesquisa, verificou-se que o algoritmo utilizado foi capaz de identificar a diferença entre os sexos, motivando a separação dos grupos pelo sexo.
- d) A RNA dedicada ao sexo masculino apresentou a seguinte arquitetura: 29 neurônios na camada de entrada e 20 neurônios na camada de saída. Assim, essa rede atingiu a eficiência global de 96,7%.

- e) A RNA dedicada ao sexo feminino apresentou a seguinte arquitetura: 18 neurônios na camada de entrada e 15 neurônios na camada de saída. Por meio dessa estrutura, essa rede também atingiu a eficiência global de 96,7%.
- f) Dessa forma, constata-se que as vozes, tanto masculinas quanto femininas, apresentam padrões que podem diferenciar uma pessoa saudável de um indivíduo com insuficiência cardíaca. Tais padrões são extraídos por meio de técnicas de análise de sinais e sua classificação pode ser realizada via redes neurais.
- g) O sistema desenvolvido apresenta uma interface gráfica de fácil acesso e uso para todos os usuários. Na plataforma gerada, alguns dados são apresentados em tabelas de forma a possibilitar a comparação entre os dados dos indivíduos saudáveis com os dos que são acometidos com IC. Além disso, gráficos comparativos também são mostrados para facilitar a análise do profissional da saúde. Ademais, o sistema fornece um possível diagnóstico ao médico, com seu grau de certeza expresso em percentual.
- h) O aplicativo elaborado apresentou os seguintes resultados: Acurácia de 91,86 %; sensibilidade de 88,1%; e especificidade de 92,1%.
- i) Assim, o presente projeto contribui para a automatização do diagnóstico da insuficiência cardíaca, possibilitando um processo mais rápido e preciso, o que supre uma das grandes dificuldades atuais do enfrentamento à IC: sua identificação em estágio ainda inicial.
- j) O sistema de identificação de insuficiência cardíaca por meio de análise sonora da voz e inteligência artificial teve como resultado final o depósito da patente sob o número BR 10 2020 017076 7 no Instituto Nacional da Propriedade Industrial (INPI).

6.2 SUGESTÕES PARA TRABALHOS FUTUROS

Para pesquisas futuras, o estudo de doenças que dificultem o diagnóstico da IC realizado pelo médico, como a pneumopatia, também podem ser adicionadas no sistema, de modo que o programa não faça apenas a diferenciação entre pessoas supostamente saudáveis daquelas com IC,

mas que o software seja capaz de auxiliar o profissional da saúde a identificar diferentes doenças, com sintomas semelhantes, de forma automática.

REFERÊNCIAS BIBLIOGRÁFICAS

AIRES, M. D. M., 2018, **“Fisiologia”**., 5ª Edição, Rio de Janeiro – RJ, Brasil, Editora: Guanabara Koogan LTDA.

ALMEIDA, A. J. S., 2010, **“Análise de Sinais Pulmonares Utilizando Técnicas no Domínio Tempo-Frequência e Classificação Neural”**., Dissertação de Mestrado em Modelagem Computacional. Universidade Federal de Alagoas, Brasil.

ACHARYA. U. R. et. al., 2008, **“Automatic Identification of Cardiac Health Using Modeling Techniques: A Comparative Study”**., Elsevier Journal.

BECKER R. O., 2018. **“Anatomia Humana”**. São Paulo – SP, Brasil. Editora: SAGAH EDUCAÇÃO S.A.

BONALDI, E. L. et al., 2014, **“Aplicação de FFT para a Detecção de Defeitos em Máquinas Elétricas”**., Congresso Nacional de Matemática Aplicada à Indústria.

BRASIL. Ministério da Saúde. Comissão Nacional de Incorporação de Tecnologia do SUS. **Peptídeos Natriuréticos tipo B (BNP e NT-ProBNP) para o diagnóstico de Insuficiência Cardíaca**. Brasília, DF, 2018.

BRASIL. Sociedade brasileira de cardiologia. Departamento de Insuficiência Cardíaca. **“Diretriz Brasileira de Insuficiência Cardíaca Crônica e Aguda”**. Brasília, DF, 2018

CAMARGO, T. A. B. S, 2011. **“Transformada Wavelet Aplicada a Análise de Falhas em Rolamentos”**. Dissertação – Universidade Estadual de Campinas Faculdade de Engenharia Mecânica, Brasil.

CANNON, C. P.; STEINBERG, B. A., 2012, **“Cardiologia Baseada em Evidências”**. 3ª Edição, São Paulo, SP – Brasil, Editora: Artmed.

Center for Disease Control and Prevention. **“Heart Failure”**. 2020. Disponível em: < https://www.cdc.gov/heartdisease/heart_failure.htm>. Acesso em 24 de março de 2021.

CERÓN I. F. C.; BADILLO A. G. G., 2011. **“A Keyword Based Interactive Speech Recognition System for Embedded Applications”**. Dissertação — Malardalen University Sweden, Suécia.

CHEN, Y., et. al., 2019. **“Feature Extraction Using Dominant Frequency Bands and Time-Frequency Image Analysis for Chatter Detection in Milling”**., Elsevier Journal.

DA SILVA, P. C. L. et. al., 2019. **“Daubechies Wavelets as Basis Functions for the Vectorial Beam Propagation Method”**., Journal of Electromagnetic Waves and Applications.

DALIMAN, S. et. al., 2017. **“Development of GUI for Automated Oil Palm Tree Counting Based on Remote Sensing Imagery”**., Institute of Electrical and Electronics Engineers (IEEE).

DAUMÉ II H., 2012. **“A Course in Machine Learning”**. Edição única. Disponível em: < ciml.info/dl/v0_99/ciml-v0_99-all.pdf > Acesso em: 25/03/2021.

DESAI A. S.; STEVENSON L.W., 2012. **“Rehospitalization for Heart Failure Predict or Prevent?”** AHA Journal

Europa. European Society of cardiology. Heart failure association. **“2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure”**. European Heart Journal, 2016.

FACELI, K.; et. al., 2011, **“Inteligência artificial: uma abordagem de aprendizado de máquina”**, Rio de Janeiro – RJ, Brasil. Editora: LTC.

GAO B.; PAVEL L, 2018. **“On the Properties of the Softmax Function with Application in Game Theory and Reinforcement Learning”**. University of Toronto, Canada.

GILS, M. V. et. al., 1997, **“Using Artificial Neural Networks for Classifying ICU Patient States”**., Institute of Electrical and Electronics Engineers (IEEE), In: Medicine and Biology.

GIUSTINO, G. et. al., 2020, **“NYHA Functional Classification and Outcomes After Transcatheter Mitral Valve Repair in Heart Failure”**., American college of cardiology foundation. Elsevier Journal.

GONCALVES, R. M. et. al., **“Modelagem preditiva de linha de costa utilizando redes neurais artificiais”**. Bol. Ciênc. Geod. [online]. vol.16, n.3, pp.420-444.

GRAAFF, V. D., 2002, **“Anatomia Humana”**., 6ª Edição, Barueri – SP, Brasil., Editora: Manole Ltda.

HAYKIN S., 2008. **“Redes Neurais – Princípios e Práticas”**. 2ª Edição, Ontário – Canadá, Bookman.

HAYKIN S., 2009. **“Neural Networks and Learning Machines”**. 3ª Edição, Ontário – Canadá, Editora: Pearson Prentice Hall.

HEMANTH, D. J. et. al., 2018, **“Analysis of Daubechies Wavelet Transform Based Human Detection Approaches in Digital Videos”**., 22nd IEEE International Conference on Intelligent Engineering Systems

INMAN, D. J., 2008, **“Engineering Vibration”**., 3rd Edition. New Jersey – NY, United States of América., Publisher: Pearson Prentice Hall.

ITTICHAICHAREON C. et al., 2012. **“Speech Recognition using MFCC”**. International Conference on Computer Graphics, Simulation and Modeling.

KARAYE, I. A. et. al., 2014, **“Analysis of Cardiac Beats Using Higher Order Spectra”**., Institute of Electrical and Electronics Engineers (IEEE).

KONAR, P.; CHATTOPADHYAY, P., 2014, **“Multi-Class Fault Diagnosis of Induction Motor Using Hilbert and Wavelet Transform”**., Elsevier Journal.

KRIESEL D., 2005. **“A Brief Introduction to Neural Networks”**. Disponível em: < <http://www.dkriesel.com> >. Acessado em: 25/03/2021.

LATHI, B. P., 2007, **“Sinais e Sistemas Lineares”**., 2ª edição, Porto Alegre – RS, Brasil., Editora: Bookman.

LEE, H. et. al., 2017. " **Impact of a Telehealth Program With Voice Recognition Technology in Patients With Chronic Heart Failure: Feasibility Study**". JMIR MHEALTH AND UHEALTH.

LIMA, T. L. d. V., 2017. **"Desenvolvimento De Um Sistema Embarcado Para Detecção De Falhas Em Veículos Automotivos Por Meio De Análise Sonora"**. Tese (Relatório de Qualificação de Doutorado) — Universidade Federal da Paraíba, Brasil.

LOSCALZO, J., 2014. **“Medicina Cardiovascular”**. 2ª edição, São Paulo – SP, Brasil., Editora: AMGH Ltda.

LYRA, G. J., 2019. **“Sistema Inteligente para Diagnóstico de Falhas em Rolamentos de Motores de Indução Trifásicos via Análise Sonora”**., Tese de Doutorado em Engenharia Mecânica. Universidade Federal da Paraíba, Brasil.

MALLAT, S., 2009. **“A Wavelet Tour of Signal Processing”**., 3rd Edition, New York – NY, United States of America., Publisher: Elsevier Inc.

MCCULLOCH, W.S., PITTS, W., 1943. **“A logical calculus of the ideas immanent in nervous activity”**. Bulletin of Mathematical Biophysics 5, 115–133 <https://doi.org/10.1007/BF02478259>

MURTON, O. M; et. al., 2017. **“Acoustic Speech Analysis of Patients with Decompensated Heart Failure: A Pilot Study”**., The Journal of the Acoustical Society of America.

NOBRE, F., 2016, **“Cardiologia de consultório: soluções práticas na rotina do cardiologista”**, 2ª edição, São Paulo – SP, Brasil, Editora: Manole.

Organização Pan-Americana da Saúde. **Doenças cardiovasculares**. Folha informativa, 2017. Disponível em: <<https://www.paho.org/pt/topicos/doencas-cardiovasculares>>. Acesso em 24 de março de 2021.

PARK, J. B., 2017, **“Impact of a Telehealth Program With Voice Recognition Technology in Patients With Chronic Heart Failure: Feasibility Study”**., US National Library of Medicine – Journal of Medical Internet Research.

PATTERSON, J.; Gibson, A., 2017, **“Deep Learning”**, 1ª edição, Sebastopol – CA, Estados Unidos da América., Editora: O’Reilly.

PENG Z. K.; TSE P. W.; CHU F.L., 2005, **“A comparison study of improved Hilbert–Huang transform and wavelet transform: Application to fault diagnosis for rolling bearing”**. Elsevier Journal.

PONIKOWSKI P. et. al., 2016, **“2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure”**, European Heart Journal.

RAY, M.; CHANDRA, M., 2017, **“Evaluation of Wavelet-Based Speech Codecs for VoIP Applications”**., Proceedings of the International Conference on Nano-electronics, Circuits & Communication Systems, Lecture Notes in Electrical Engineering, Chapter 3, Singapura.

REDDY, M. K. et. al.; 2021. **“The automatic detection of heart failure using speech signals”**. Elsevier Journal, Computer Speech & Language.

RESNIK R. R.; MISCH C. E., 2017., **“Avoiding Complications in Oral Implantology”**. Editora: Mosby. ISBN: 0323375804.

RIBEIRO G. L. et. Al., 2014. **“Análise mel-cepstral na discriminação de patologias laríngeas”**. XXIV Congresso Brasileiro de Engenharia Biomédica.

RIELLA, R. J., 2008, **“Propostas de Técnicas para Caracterização e Classificação Automática de Sons Pulmonares Adventícios”**., Tese de Doutorado em Ciências. Universidade Tecnológica Federal do Paraná, Brasil.

RIZZO, D. C., 2012, **“Fundamentos da Anatomia e Fisiologia”**., 3ª edição, São Paulo – SP, Brasil., Editora: Cengage Learning.

ROBERTS, M. J., 2009, **“Fundamentos em Sinais e Sistemas”**., São Paulo – SP, Brasil., Editora: AMGH Ltda.

ROSSETO C. et. al., 2019, **“Causas de internação hospitalar e óbito em idosos brasileiros entre 2005 e 2015”**. Revista Gaúcha de Enfermagem. vol.40 Porto Alegre 2019.

SANTOS N. C. M, 2014, **“Anatomia e fisiologia humana”**. São Paulo – SP, Brasil., Editora: Érica ISBN: 978-85-365-1095-8

Secretaria de saúde do estado de Goiás. **Dia do Coração alerta população para prevenção das doenças cardíacas**. Dia do Coração alerta população para prevenção das doenças cardíacas, 2019. Disponível em: < <https://www.saude.go.gov.br/noticias/132>>. Acesso em 24 de março de 2021.

SEO, Y., et. al., 2018, **“Drone Detection Using Convolutional Neural Networks with Acoustic STFT Features”**., Institute of Electrical and Electronics Engineers (IEEE).

SHALEV-SHWARTZ, S.; BEN-DAVID, S., 2014. **“Understanding Machine Learning: From Theory to Algorithms**. Cambridge: Cambridge University Press. doi:10.1017/CBO9781107298019.

SHERWOOD, L., 2011., **“Fisiologia Humana”**., 7ª edição, São Paulo – SP, Brasil., Editora: Cengage Learning.

UZINSKI, J. C., 2013, **“Momentos Nulos e Regularidade Wavelet na Detecção de Falhas em Sinais”**., Dissertação de Mestrado em Engenharia Elétrica, Universidade Estadual Paulista, Brasil.

WILAMOWSKI B. M.; IRWIN J. D, 2011. **“Intelligent systems”**. 2ª edição. Editora: The Electrical Engineering Handbook Series.

WINURSITO A. et. al., 2018. **“Improvement of MFCC Feature Extraction Accuracy Using PCA in Indonesian Speech Recognition”**. International Conference on Information and Communications Technology.

YAN, R. et. al., 2013, **“Wavelets for Fault Diagnosis of Rotary Machines: A Review with Applications”**., Elsevier Journal.