

Bruno Nogueira de Carvalho

**DIAGNÓSTICO AUTOMÁTICO DE  
PATOLOGIAS CARDÍACAS BASEADO EM  
DESCRITORES SONOROS E SVM**

São Luis - MA

2018

Bruno Nogueira de Carvalho

**DIAGNÓSTICO AUTOMÁTICO DE PATOLOGIAS  
CARDÍACAS BASEADO EM DESCRITORES  
SONOROS E SVM**

Monografia apresentada ao curso de Ciência da Computação da Universidade Federal do Maranhão, **como parte dos requisitos necessários** para obtenção do grau de Bacharel em Ciência da Computação.

Universidade Federal do Maranhão – UFMA

Curso de Ciência da Computação

Orientador: Prof. Dr. João Dallyson Sousa de Almeida

São Luis - MA

2018

Ficha gerada por meio do SIGAA/Biblioteca com dados fornecidos pelo(a) autor(a).  
Núcleo Integrado de Bibliotecas/UFMA

Carvalho, Bruno Nogueira de.

Diagnóstico automático de patologias cardíacas baseado em descritores sonoros e SVM / Bruno Nogueira de Carvalho. - 2018.

52 f.

Orientador(a): João Dallyson Sousa de Almeida.

Monografia (Graduação) - Curso de Ciência da Computação, Universidade Federal do Maranhão, São Luis, 2018.

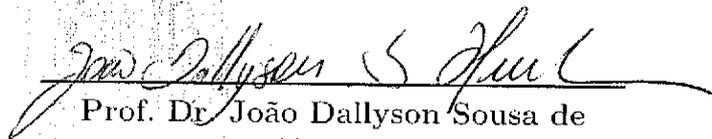
1. Doenças cardíacas. 2. Máquina de Vetores de Suporte. 3. SMOTE. 4. Sons cardíacos. 5. Transformada wavelet. I. Almeida, João Dallyson Sousa de. II. Título.

Bruno Nogueira de Carvalho

# DIAGNÓSTICO AUTOMÁTICO DE PATOLOGIAS CARDÍACAS BASEADO EM DESCRITORES SONOROS E SVM

Monografia apresentada ao curso de Ciência da Computação da Universidade Federal do Maranhão, **como parte dos requisitos necessários** para obtenção do grau de Bacharel em Ciência da Computação.

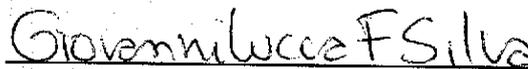
São Luis - MA, 29 de novembro de 2018:



Prof. Dr. João Dallyson Sousa de  
Almeida  
Orientador



Prof. Dr. Tiago Bonini Borchardt  
Convidado 1



Prof. Me. Giovanni Lucca França da  
Silva  
Convidado 2

São Luis - MA

2018

*Aos meus pais, irmã, meus amigos e meu orientador que, com muito carinho e apoio, me ajudaram a escalar essa montanha celeste e não mediram esforços para que eu chegasse até esta etapa de minha vida.*

# Agradecimentos

Agradeço, principalmente, à minha família que amo tanto, minha mãe Marise, meu pai Ivaldo e minha irmã Cintia, por serem de fundamental importância no meu crescimento e desenvolvimento pessoal, por me prepararem todos os dias para o que eu poderia enfrentar fora da nossa casa, por me darem a oportunidade de crescer intelectualmente me colocando ao alcance da informação e do conhecimento que adquiri ao longo dos anos e por tudo que fazem todos os dias para me ajudar a me tornar uma pessoa melhor.

Devo também grande parte da minha formação a todos os meus professores e colegas desde a escola até o final do curso de graduação que de alguma maneira não apenas repassaram seus conhecimentos, mas me incentivaram a pensar, a desenvolver outras sensibilidades e a me preparar para a vida. Ao meu orientador e professor João Dallyson Almeida e a professora e coordenadora Simara Vieira da Rocha por me guiarem durante o período dessa monografia, apontando meus erros e que direção deveria seguir.

Agradeço também aos meus colegas de faculdade que foram essenciais na formação de opiniões e trocas de conhecimento durante o curso, em especial a Thalles Alencar, Italo Francyles e Maurício César, que nunca teria conseguido chegar até aqui sem a ajuda deles.

*O que realmente importa na vida é o que  
se faz com o tempo que nos é dado.*

*J. R. R. Tolkien*

# Resumo

Todos os anos milhões de pessoas morrem em decorrência de doenças cardíacas. A auscultação é um dos exames médicos mais simples e populares para diagnosticar a presença de anomalias cardíacas e respiratórias. O objetivo do presente estudo é apresentar um método adequado que possa detectar e classificar sons de batimentos cardíacos com algum tipo de anormalidade. A relevância do estudo está em apresentar meios que suscitem maiores estudos no futuro que possam levar ao desenvolvimento de uma ferramenta de diagnóstico inteligente. As amostras obtidas de sons cardíacos são filtradas através de uma transformada *wavelet* e tem suas características extraídas por seis descritores de áudio diferentes (MFCC, *Mel Spectrogram*, características *chroma*, *spectral contrast*, *tonnetz* e *zero-crossing rate*). Os dados extraídos desses descritores são inseridos num classificador máquina de vetores de suporte. Através dessa modelagem, a metodologia proposta conseguiu alcançar 78,32% e 84,1% de acurácias nas duas bases de dados utilizadas.

**Palavras-chave:** Doenças cardíacas; Sons cardíacos; Transformada wavelet; Máquina de Vetores de Suporte; SMOTE;

# Abstract

Every year millions of people die from heart disease. Auscultation is one of the simplest and most popular medical tests to diagnose the presence of cardiac and respiratory anomalies. The objective of the present study is to present a suitable method that can detect and classify sounds of heart beats with some type of abnormality. The relevance of the study is to present means that will lead to further studies in the future that may lead to the development of an intelligent diagnostic tool. Samples obtained from cardiac sounds are filtered through a wavelet transform and have their characteristics extracted by six different audio descriptors (MFCC, Mel Spectrogram, chroma features, spectral contrast, tonnetz and zero crossing rate). The data extracted from these descriptors are inserted into a classifier of carrier vectors. Through this modeling, the proposed methodology was able to achieve 78.32 % and 84.1 % accuracy in the two databases used.

**Keywords:** Heart diseases; Cardiac sounds; Wavelet transform; Support Vector Machine; SMOTE;

# Lista de ilustrações

Figura 1 – Localização dos pontos de Auscultação: Ponto 1 - mancha mitral; Ponto 2 - tricuspídeo; Ponto 3 - pulmonar; Ponto 4 aórtico . . . . .	21
Figura 2 – Exemplos de sopros diastólicos (primeira e terceira divisões), sistólicos (segunda divisão) e contínuos e patologias correspondentes . . . . .	23
Figura 3 – Bloco básico de construção de Ingrid Daubechies . . . . .	24
Figura 4 – Exemplo de Hiperplano do SVM . . . . .	27
Figura 5 – Gráficos de PCG e respectivos envelopes instantâneos de amplitude e energia. . . . .	29
Figura 6 – Gravação do PCG . . . . .	30
Figura 7 – Os componentes separados de um sinal normal de PCG. De cima para baixo: ruído de fundo, componente mitral de S1, S3, componente aórtico de S2, S4, componente tricúspide de S1 e componente pulmonar de S2. . . . .	32
Figura 8 – ECG sincronizado (superior) e PCG (inferior) mostrando o tempo do complexo QRS-S1 e T-onda-S2 relações . . . . .	33
Figura 9 – Etapas do método proposto . . . . .	37
Figura 10 – Exemplos de áudios originais de <i>The PASCAL Classifying Heart Sounds Challenge 2011</i> . . . . .	39
Figura 11 – <i>Wavelet</i> Daubechies de Ordem 6 . . . . .	40

# Lista de tabelas

Tabela 1 – Exemplo de Matriz de Confusão . . . . .	27
Tabela 2 – Exemplo de geração de exemplos sintéticos. . . . .	41
Tabela 3 – Matriz de Confusão dos resultados da classificação após filtragem da Base A . . . . .	42
Tabela 4 – Precisão detalhada por classe após filtragem da base A . . . . .	42
Tabela 5 – Matriz de Confusão dos resultados da classificação após filtragem da Base B . . . . .	42
Tabela 6 – Precisão detalhada por classe após filtragem da base B . . . . .	43
Tabela 7 – Re-amostragem das instâncias da Base A. . . . .	43
Tabela 8 – Matriz de Confusão dos resultados da classificação após filtragem e SMOTE na Base A. . . . .	44
Tabela 9 – Precisão detalhada por classe após filtragem, normalização e SMOTE da base A . . . . .	44
Tabela 10 – Re-amostragem das instancias da Base B . . . . .	44
Tabela 11 – Matriz de Confusão dos resultados da classificação após filtragem e SMOTE da Base B . . . . .	44
Tabela 12 – Precisão detalhada por classe após filtragem, normalização e SMOTE da base B . . . . .	44
Tabela 13 – Resultados após Seleção de Características . . . . .	45
Tabela 14 – Comparação de resultados com Gomes e Pereira (2012), de Deng e Bentley (2012) e da monografia na Base A . . . . .	46
Tabela 15 – Comparação de resultados com Gomes e Pereira (2012), de Deng e Bentley (2012) e da monografia na Base B . . . . .	46

# Lista de abreviaturas e siglas

AVC	Acidente Vascular Cerebral
DAC	Doença Arterial Coronariana
ECG	Eletrocardiograma ( <i>Electrocardiogram</i> )
FN	Falso Negativo
FP	Falso Positivo
Hz	O hertz é a unidade de medida para frequência
MFCC	Coefficientes cepstrais frequência mel ( <i>Mel-Frequency Cepstral Coefficients</i> )
MLP	Rede neural Perceptron multicamadas ( <i>MultiLayer Perceptron</i> )
PCG	Fonocardiograma (Phonocardiogram)
QRS	O complexo QRS representa a despolarização do músculo ventricular no eletrocardiograma.
SMOTE	Técnica de Sobreposição de Minoria Sintética ( <i>Synthetic Minority Over-sampling TEchnique</i> )
SVM	Máquina de vetores de suporte ( <i>Support Vector Machine</i> )
TFR	Técnicas de Fisioterapia Respiratória
VN	Verdadeiro Negativo
VP	Verdadeiro Positivo
WAV	Formato-padrão de arquivo de áudio da Microsoft

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>13</b>
<b>1.1</b>	<b>Objetivos</b>	<b>14</b>
1.1.1	Específicos	14
<b>1.2</b>	<b>Organização do Trabalho</b>	<b>15</b>
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>16</b>
<b>2.1</b>	<b>Panorama Geral Sobre Fisiologia Cardíaca</b>	<b>16</b>
2.1.1	Anatomia do Coração	16
2.1.2	Algumas Doenças Cardíacas	17
2.1.3	Auscultação Cardíaca	19
2.1.3.1	Os Sons do Coração	20
2.1.3.1.1	Ritmos de galope	22
2.1.3.1.2	Murmúrios	22
<b>2.2</b>	<b>Técnicas computacionais</b>	<b>24</b>
2.2.1	Transformada Wavelet	24
2.2.2	Descritores Sonoros	25
2.2.3	Classificação	26
2.2.4	Métricas de Avaliação a Classificação	27
<b>3</b>	<b>TRABALHOS RELACIONADOS</b>	<b>29</b>
<b>3.1</b>	<b>Algoritmos de Classificação Cardíacos</b>	<b>34</b>
<b>4</b>	<b>MATERIAIS E MÉTODO</b>	<b>37</b>
<b>4.1</b>	<b>Descrição da Base de Dados</b>	<b>37</b>
<b>4.2</b>	<b>Rotinas de Extração de Características</b>	<b>38</b>
<b>4.3</b>	<b>Normalização e Re-amostragem de Instâncias</b>	<b>39</b>
<b>4.4</b>	<b>Classificação</b>	<b>41</b>
<b>5</b>	<b>RESULTADOS E DISCUSSÃO</b>	<b>42</b>
<b>5.1</b>	<b>Resultados obtidos com geração de exemplos sintéticos (SMOTE)</b>	<b>43</b>
<b>5.2</b>	<b>Resultados com seleção de características</b>	<b>45</b>
<b>5.3</b>	<b>Comparação de resultados</b>	<b>46</b>
<b>6</b>	<b>CONCLUSÃO</b>	<b>48</b>
	<b>REFERÊNCIAS</b>	<b>49</b>

# 1 Introdução

No ambiente clínico em constante desenvolvimento, ocasionalmente há oportunidades para certas modalidades de diagnóstico para surgir e tornar-se parte da prática clínica. A auscultação assistida por computador consiste no processamento digital de um sinal de fonocardiograma (ou PCG), que é um gráfico de gravação de alta fidelidade dos sons e sopros feitos pelo coração, para auxiliar os médicos especialistas na tarefa de interpretá-lo corretamente e elaborar o diagnóstico.

Estima-se que 17 milhões de pessoas morreram devido ao fumo e à doença arterial coronariana em 2016, sendo os maiores causadores de mortes do mundo (ORGANIZATION; ORGANIZATION et al., 2017). Dentre as doenças cardiovasculares que mais matam estão: 1) Infarto do miocárdio, também conhecido como ataque cardíaco cuja característica é a diminuição da circulação sanguínea no coração; 2) Doença vascular periférica, que obstrui as artérias periféricas do corpo, causando, por exemplo, dormência; 3) o Acidente Vascular Cerebral (AVC), também denominado de derrame e 4) Morte súbita, ocasionada por doenças congênitas cardíacas desenvolvidas com o passar dos anos sem a percepção da mesma (ORGANIZATION; ORGANIZATION et al., 2017). Cerca de 300 mil brasileiros morrem de doenças cardiovasculares por ano (BRASIL, 2018).

Tendo em vista o número de mortes registradas cuja causa são as doenças cardíacas, verifica-se a necessidade de identificar com antecedência tais doenças. Dentre as inúmeras metodologias de diagnósticos existentes, o uso da ausculta cardíaca se faz essencial para perceber os focos, bulhas, ritmo, frequência<sup>1</sup> e outras possíveis alterações que ocasionam patologias.

Entretanto, para detectar anormalidades através do som é necessário a experiência e conhecimento do examinador. Esse problema atrai interesse para classificação dos sons cardíacos normais e anormais, afim de melhor auxiliar o medico na tomada de decisão e, também, na prevenção de doenças.

Cabe revelar que a importância da auscultação correta e prévia tem sido comprovada através do uso clínico extensivo ao longo dos últimos séculos e, conseqüentemente, mostra se importante o desenvolvimento de ferramentas que auxiliem os especialistas no difícil processo de avaliação da auscultação seria mais que bem-vinda (GOMES; PEREIRA, 2012).

Uma das principais áreas no âmbito desse campo de oportunidade é a detecção

---

<sup>1</sup> Focos são onde se podem melhor auscultar fenômenos cardíacos, como bulhas; bulhas são os sons gerados no tórax pelo impacto do sangue na circulação sanguínea; ritmo cardíaco mede o padrão de batimentos cardíacos e pode identificar anomalias neles; frequência é a quantidade de batidas que o coração faz por minuto

de sopros para avaliar cardiopatias importantes. Murmúrios, sintomáticos de algumas cardiopatias, podem ser registradas em um PCG e, embora os cardiologistas possam facilmente percebê-las, clínicos gerais podem ter mais dificuldade e, portanto, um método computacional capaz de detectar murmúrios, caracterizar e diagnosticar ou auxiliar no diagnóstico de um paciente é de grande importância. Cumpre destacar que esse sistema deve ter mais especificidade, não só em termos de algoritmo, mas também nos métodos de aquisição de bases de dados e classificação, pois assim facilita o diagnóstico pelos outros profissionais da saúde (GOMES; PEREIRA, 2012).

A justificativa para a escolha do presente estudo foi pautada na observação de que 90% dos casos de doenças cardiovasculares podem ser evitados por meio de medidas de prevenção, como exames preliminares de auscultação (MCGILL; MCMAHAN; GIDDING, 2008).

## 1.1 Objetivos

Neste contexto, o presente estudo objetiva apresentar um método automático de detecção de anomalias de sons cardíacos, classificando-os com base em suas características extraídas. Para tanto serão utilizadas filtragem de sinais com transformada *wavelet*, técnicas de extração de características sonoras e musicais da biblioteca librosa (MCFEE et al., 2018) e classificação dos sons cardíacos.

### 1.1.1 Específicos

- a) Detectar anormalidades através dos sons dos batimentos cardíacos.
- b) Adquirir base de sons de doenças cardíacas. Obter características de padrão sonora.
- c) Validar os resultados.

## 1.2 Organização do Trabalho

Feitas as devidas considerações o presente estudo monográfico apresenta cinco capítulos que consubstanciarão paulatinamente o objetivo pretendido pelo presente estudo.

O primeiro capítulo é composto da Introdução que traz elementos essenciais à pesquisa, apontando a justificativa para a escolha do tema, o objetivo do trabalho, bem como expõe um breve resumo de como os capítulos serão desenvolvidos.

O segundo capítulo apresenta um panorama geral sobre a Fisiologia cardíaca, e serão feitas considerações sobre o coração, a anatomia deste e algumas doenças apresentadas.

Feitas as devidas considerações caberá ao quinto capítulo expor sobre a análise dos resultados que foram apresentados; também neste capítulo será feita a discussão pautada nos resultados obtidos. Por fim, o sexto capítulo apresenta, a seção Conclusão sobre os resultados e trabalhos futuros.

## 2 Fundamentação Teórica

Neste capítulo apresenta-se as principais características do coração, fazendo compreender que para que haja a percepção de que há algum sinal de distúrbio cardíaco, se faz necessário antes apresentar de que forma o coração se comporta, os batimentos, os sons dentre outras características cardíacas.

Descrevem-se, ainda, conceitos computacionais importantes para a compreensão do método proposto neste trabalho.

### 2.1 Panorama Geral Sobre Fisiologia Cardíaca

#### 2.1.1 Anatomia do Coração

O coração é um órgão muscular que bombeia o sangue através de seu sistema circulatório (o sistema de transporte de sangue que inclui o coração, artérias, capilares e veias) para todas as partes do corpo.

Está dividido em quatro câmaras. As duas câmaras superiores, chamadas átrios, são unidas às duas câmaras inferiores, chamadas de ventrículos. Uma partição interna separa o átrio e ventrículo esquerdo do átrio e ventrículo direito.

Os átrios recebem sangue que está retornando ao coração da circulação, que então flui para os ventrículos - as bombas do coração. A bomba ventricular direita envia sangue para fora do coração para os pulmões, enquanto a bomba ventricular esquerda envia sangue para o resto de seus tecidos e órgãos. A parede do músculo ventricular esquerdo é mais espessa do que a direita, pois bombear o sangue para os pulmões é uma tarefa relativamente fácil em comparação com bombear todo o corpo ([AMIRI; ARMANO, 2013](#)).

O sangue circula pelas câmaras do coração em uma sequência específica. O ventrículo direito bombeia sangue do coração para os pulmões, onde ele capta oxigênio. Ele retorna dos pulmões e flui para o átrio esquerdo e depois para o ventrículo esquerdo, que o bombeia de volta para o sistema circulatório que supre o resto do seu corpo, e finalmente retorna ao coração, fluindo para o átrio direito. Válvulas unidirecionais entre as câmaras do coração mantêm o sangue circulando pelo coração na direção certa. As válvulas têm abas, chamadas cúspides, que se abrem quando o sangue é bombeado através delas e se fecham para evitar o refluxo.

Os sons do coração são produzidos quando as válvulas atrioventriculares, seguidas pelas válvulas semilunares, se fecham rapidamente. A ciclagem de sangue pelo coração ocorre em duas fases: a diástole, o período de tempo em que os ventrículos ficam relaxados,

permitindo que o sangue flua para eles a partir dos átrios; e sístole, o período de tempo em que os ventrículos se contraem e ejetam sangue para a aorta e para a veia pulmonar. Essas fases formam um ciclo cardíaco ou batimento cardíaco (GEDDES, 2005).

Os médicos usam uma ou mais medidas físicas para avaliar o quanto o coração está funcionando. Estes incluem o débito cardíaco, que é o volume de sangue que o coração bombeia através do sistema circulatório em um minuto; volume sistólico, que é o volume de sangue bombeado pelo ventrículo esquerdo em um batimento cardíaco; e a fração de ejeção, que é a fração de sangue bombeada do coração pelos seus ventrículos a cada batida do coração. Caso isto ocorra com um adulto saudável, a faixa normal para o débito cardíaco é de 4 l/min a 8 l/min, seu volume de ejeção estará entre 55 ml e 100 ml e a fração de ejeção entre 55% e 70% (AMIRI; ARMANO, 2013).

### 2.1.2 Algumas Doenças Cardíacas

O coração é como qualquer outro músculo do corpo. Ele precisa de um suprimento sanguíneo adequado para fornecer oxigênio, para que o músculo possa contrair e bombear o sangue para o resto do corpo. O coração não apenas bombeia o sangue para o resto do corpo, mas também bombeia o sangue para si mesmo através das artérias coronárias. Essas artérias são originárias da base da aorta (o principal vaso sanguíneo que transporta o sangue oxigenado do coração) e, em seguida, ramificam-se ao longo da superfície do coração (NASERI; HOMAEINEZHAD, 2013).

Quando houver o estreitamento de uma ou mais artérias, isso pode dificultar a chegada de sangue adequado ao coração, especialmente durante o algum tipo de atividade física. Assim, se as artérias continuarem a estreitar-se haverá maior dificuldade na chegada de sangue, isso pode provocar dores no músculo cardíaco, e nesse sentido, é melhor que haja diminuição na intensidade das atividades físicas para não haver estresse no coração. Os sintomas clássicos de dor torácica ou pressão e falta de ar que geralmente se espalham para os ombros, braços e/ou pescoço devido à doença cardíaca aterosclerótica (ASHD) ou doença arterial coronariana (DAC) são chamados de angina (NASERI; HOMAEINEZHAD, 2013).

Se uma das artérias coronárias ficar completamente bloqueada - geralmente devido a uma placa que se rompe e provoca a formação de um coágulo sanguíneo - o suprimento de sangue para uma parte do coração pode se perder. Isso faz com que um pedaço do músculo cardíaco morra. Isso é chamado de ataque cardíaco ou infarto do miocárdio (mio = músculo + cárdia = coração + infarto = morte do tecido) (GOMES; JORGE; AZEVEDO, 2013).

A insuficiência cardíaca é causada por qualquer condição que danifique ou enfraqueça o coração, reduzindo sua capacidade de bombear sangue com eficiência, como

ataques cardíacos, pressão alta, doença arterial coronariana ou infecção. Inicialmente, o coração tenta compensar a perda do débito cardíaco, desenvolvendo mais massa muscular, aumentando e bombeando mais rápido. As alterações também podem ocorrer no sistema circulatório, pois o corpo se ajusta para diminuir o débito cardíaco. Por exemplo, os vasos sanguíneos podem se estreitar para aumentar a pressão sanguínea, e o sangue pode ser desviado de tecidos e órgãos menos essenciais, como os rins, enquanto o corpo tenta compensar o poder reduzido do coração (MCGILL; MCMAHAN; GIDDING, 2008).

Com o passar do tempo, o aumento da carga de trabalho leva a alterações no próprio músculo cardíaco (conhecido como remodelação), fazendo com que ele endureça, reduzindo ainda mais sua capacidade de bombear o sangue eficientemente, o que leva ao agravamento da insuficiência cardíaca. Conforme o coração se esforça cada vez mais para acompanhar a demanda do corpo por sangue e oxigênio, sinais e sintomas característicos de insuficiência cardíaca começam a aparecer (TAVEL, 2006).

A redução do débito cardíaco e suprimento de sangue podem ter múltiplos efeitos em vários órgãos e tecidos. Esses incluem:

- Frequência cardíaca aumentada devido ao aumento da atividade do sistema cardiovascular, a parte do sistema nervoso responsável pela aceleração da frequência cardíaca, constrição dos vasos sanguíneos e elevação da pressão arterial. Embora o aumento da frequência cardíaca ajude a manter o débito cardíaco, esse aumento também requer mais oxigênio para o coração – nesse sentido o aumento da perfusão <sup>1</sup> cardíaca pode piorar a doença arterial coronariana ou causar arritmias, além de estimular o aumento da massa muscular levando ao aumento do coração.
- Aumento da pressão arterial devido à atividade do sistema nervoso simpático (aumenta a quantidade de trabalho que o coração deve realizar).
- Aumento do volume sanguíneo e da pressão arterial devido à secreção do hormônio antidiurético em resposta à atividade do sistema nervoso simpático, que causa retenção de líquidos nos rins.
- Sal e retenção de líquidos pelos rins causado como resultado da redução do fluxo sanguíneo para os rins, o que leva à secreção de aldosterona, um hormônio que estimula a absorção de sódio pelos rins, e regula o equilíbrio de sal e água na corrente sanguínea.
- Remodelação do músculo cardíaco causada por níveis cronicamente altos de vários hormônios, incluindo catecolaminas, renina, angiotensina e aldosterona.

<sup>1</sup> A perfusão é a medida do volume de sangue que flui através dos capilares em um tecido. Essa medida faz compreender a saúde de um tecido particular, uma vez que a perfusão ineficaz significa que o fluxo sanguíneo não está circulando adequadamente, por exemplo não está levando nutrientes adequados às células do tecido em questão.

- Diminuição da força muscular devido à atrofia do músculo esquelético resultante da redução da perfusão.
- Função hepática prejudicada e icterícia causada por congestão hepática grave

A insuficiência cardíaca é uma síndrome clínica (um distúrbio que causa um grupo de sinais e sintomas) e não uma doença específica. A combinação de sintomas que se experimenta dependerá de quaisquer condições subjacentes que se possa ter, bem como os maus funcionamentos específicos dentro do próprio coração. Os sintomas mais comuns incluem falta de ar, especialmente com exercícios ou quando ao deitar, fadiga, fraqueza e inchaço nas pernas, tornozelos e pés.

Embora todos esses sinais e sintomas possam ser causados por insuficiência cardíaca, eles também ocorrem comumente como resultado de outras doenças cardíacas ou pulmonares; por exemplo, arritmia ou embolia pulmonar. A combinação de sintomas depende em grande parte de que lado do peito está doendo, por exemplo, sentir dor no lado esquerdo do peito, irradiando para o braço, é sintoma clássico de infarto no miocárdio; já as dores no lado direito, podem ser inflamação nos órgãos, como a pericardite – inflamação na membrana que envolve o coração (HANNA; SILVERMAN, 2002).

### 2.1.3 Auscultação Cardíaca

A prática da auscultação foi registrada pela primeira vez durante o período hipocrático.<sup>2</sup> Nesse período era feita através da aplicação da orelha no peito ou no abdômen para ouvir os sons de dentro do corpo. Esta abordagem é chamada de auscultação imediata porque não usa nenhum aparelho para transportar som do corpo para o médico. Hipócrates descreveu vários sons internos diferentes a partir do qual diferentes diagnósticos poderiam ser formulados como: "Você deve saber com isto que o peito contém água, mas não pus, se na aplicação do ouvido durante um certo tempo no lado, você percebe um barulho como o de vinagre em ebulição (GOMES; JORGE; AZEVEDO, 2013).

Uma descrição formal dos sons do coração, no entanto, só foi concluída em 1628 em William Harvey's *De Motu Cordis*, no qual ele conclui que a principal função do coração é bombear sangue para as extremidades do corpo. Para caracterizar os sons do coração e seu reconhecimento adequado, é essencial que médico e paciente estejam em um ambiente adequado (sem ruídos externos). Em geral a ausculta é realizada com o paciente em decúbito dorsal são caracterizados característica auscultatória de B1, que é um som mais suave e prolongado, como uma batida de surdo ("Tum"), e de B2, que é um som mais seco ("Tá"). (KUMAR et al., 2008).

<sup>2</sup> A medicina era exercida por sacerdotes-médicos, que interpretavam sonhos, e os tratamentos das doenças era feito através de banhos, massagens, dietas, dentre outros meios. Assim Hipócrates foi o primeiro filósofo que deixou de lado a crença de que as doenças eram causadas por maus espíritos. Assim Hipócrates é considerado o pai da medicina.

De maneira geral, a ausculta continuou sendo uma técnica usada por um grupo seleto de profissionais, sendo a palpação e a percussão o exame principal e as técnicas utilizadas nesse período.

Em 1819, houve sensível alteração nessa situação. Foi, quando René Théophile Hyacinthe Laennec inventou ausculta mediada. Diante dos problemas de ausculta imediata, Laennec usou um rolo de papel rolo em um cilindro para ouvir os sons do coração de um paciente. Mais tarde, ele desenvolveu um aparelho adequado constituído por um cilindro de madeira perfurado, um tampão em forma de funil e uma rolha. Ele nomeou esse dispositivo como Cilindro ou Estetoscópio (do grego *stēthos* 'peito' + *skopein* 'exame') (HANNA; SILVERMAN, 2002).

Em pouco tempo, o estetoscópio ganhou popularidade e apesar da falta de exatidão de suas interpretações dos sons do coração, na década de 1830 era uma ferramenta de cabeceira inegável para o exame de problemas no peito e algo esperado pelos pacientes que consultam um médico. O uso do estetoscópio cresceu ao longo dos séculos 19 e 20 para se tornar uma ferramenta médica de excelência que poderia obter resultados notáveis, dependendo de um treinamento médico para o uso correto (HANNA; SILVERMAN, 2002).

Estetoscópios mais recentes receberam a capacidade de gravar sons e transmiti-los a um computador não só para uma revisão posterior e acompanhamento do paciente, mas também para processamento de computador e análise. Essas novas modalidades trazem maiores custos para os hospitais devido os valores elevados de equipamentos e por que ainda não se encontram com tanta facilidade no mercado. Outra situação acerca do uso desses novos estetoscópios trata-se do um atraso no diagnóstico final do paciente, uma vez que os profissionais ainda não estão adaptados ao uso dessas tecnologias o que resulta em maior sofrimento na esperas, à medida que aguardam o diagnóstico e testes sequenciais. Em consequência, recentemente foi dada atenção aos sistemas de apoio à decisão para ajudar na primeira triagem clínicos gerais tomar decisões mais rápidas e mais baratas usando uma ferramenta que eles são certo para levar junto com eles em todos os momentos, o estetoscópio (TAVEL, 2006).

### 2.1.3.1 Os Sons do Coração

O coração é o centro do sistema cardiovascular. Serve como uma bomba com o propósito de transportar o sangue dos tecidos para os pulmões e vice-versa. Para executar esta ação, o coração, passa por uma sequência de eventos de forma organizada e precisa. Cada sequência de eventos do coração, desde o início de um batimento cardíaco até o começo do próximo é chamado de ciclo cardíaco (GHOSH, 2010).

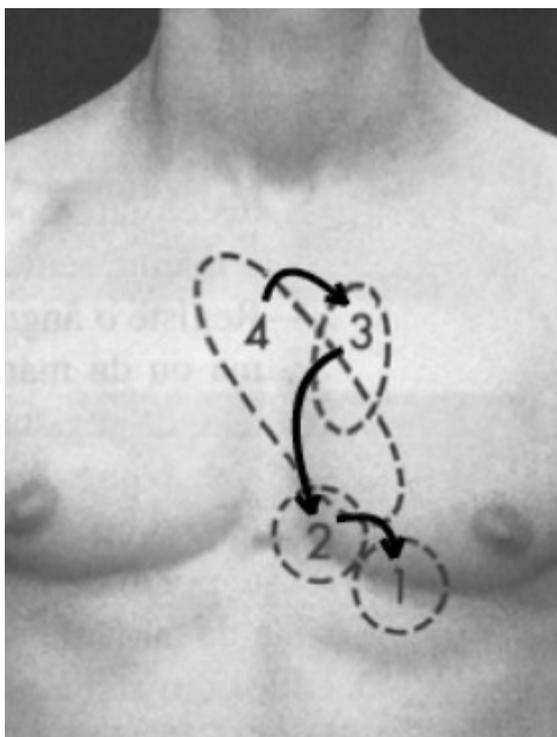
Os sons do coração são gerados pelos batimentos cardíacos e pelo fluxo sanguíneo. Sempre que ocorre o fluxo sanguíneo turbulento, este provoca vibrações que podem ser ouvidas pelo ouvido humano. Fluxo sanguíneo laminar, que ocorre em quase todo o espectro

do sistema cardiovascular, e este não produz som. Por muito tempo pensou-se que as próprias vibrações causadas pelo fechamento das válvulas produzia os sons do coração, mas esta teoria foi provada errada (BENTLEY et al., 2011).

O coração de um adulto saudável produz dois sons (S1 e S2), frequentemente descritos como primeira e segunda bulha <sup>3</sup> cardíaca. Existem também os sons adicionais também denominados de terceira ou quarta bulha, como sopros e ritmos de galope (S3 e S4) também podem estar presentes e podem ou não ser um sinal de patologia.

Em indivíduos adultos, existem dois sons a primeira bulha e a segunda bulha que geralmente são considerados sons normais, já a terceira bulha e a quarta bulha, que só apresentam sons normais em algumas circunstâncias. Os sons que ocorrem em cada batida do coração são propagados pelo tórax e são irradiados para diferentes áreas do peito, dependendo da origem, normalmente a auscultação é feita nas chamadas áreas ou pontos de auscultação: a aorta, pulmonar, tricúspide e mitral. A localização desses pontos é mostrada na Figura 1 e sua faixa de frequência total se estende dentro de 20-1000Hz.

Figura 1 – Localização dos pontos de Auscultação: Ponto 1 - mancha mitral; Ponto 2 - tricuspídeo; Ponto 3 - pulmonar; Ponto 4 - aórtico



Fonte: (DELGADO-TREJOS et al., 2009)

Os primeiro e segundo sons do coração são completamente não-patológicos e correspondem a turbulência do sangue causada pelo fechamento das válvulas cardíacas. S1

<sup>3</sup> Bulhas cardíacas são sons gerados pelo impacto do sangue em diversas estruturas cardíacas . As vibrações são propagadas às paredes do tórax e podem ser auscultadas através do estetoscópio.

é causado pelo fechamento das Válvulas atrioventriculares e S2 é causado pelo fechamento das válvulas semilunares.

Esses dois sons pode então ser usado como limites entre a sístole e a diástole. A sístole ocorre entre S1 e S2 e diástole de S2 para S1. S1 e S2 são normalmente os eventos de maior amplitude em um PCG e têm frequências entre 20-200Hz. A duração normal desses sons varia de 70ms a 140ms. O seu espectro de frequências é muito semelhante, mas foi demonstrado que os S2s têm amplitudes maiores acima de 150Hz do que os S1s (NASERI; HOMAEINEZHAD, 2013).

#### 2.1.3.1.1 Ritmos de galope

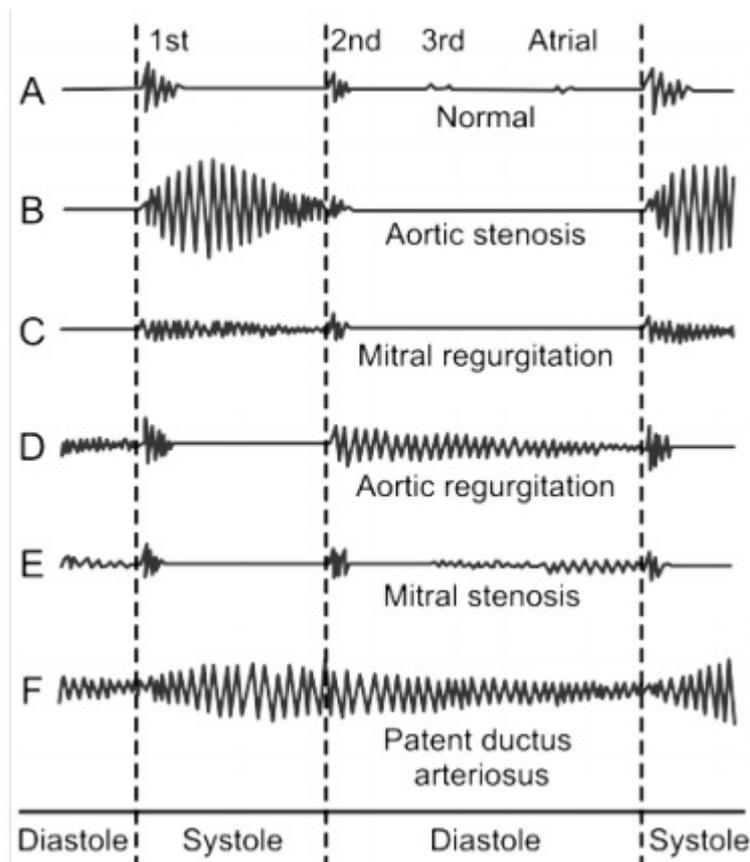
Além de S1 e S2, os ritmos de galope contém os sons extras S3 e / ou S4. O termo ritmos de galope vem do fato de que, em vez da sequência usual de som, um som adicional é ouvido causando um som similar a um galope. Ambos os sons ocorrem dentro da diástole, mas S3 vem logo após S2, enquanto S4 vem logo antes de S1 (HANNA; SILVERMAN, 2002).

A origem do S3 é discutida mas acredita-se que seja causado pela tensão das cordas, enquanto o sangue distende o ventrículo esquerdo. Isto ocorre normalmente em crianças e adultos jovens, mas é, em outros assuntos, geralmente patologicamente indicando insuficiência cardíaca (BENTLEY et al., 2011). O quarto som cardíaco, o S4 é muitas vezes audível em crianças saudáveis, quando é auscultado em adultos é denominado de galope pré-sistólico, e este som de galope, é um sinal de estado patológico. O S4 é conhecido por ser causado pela força do sangue do átrio para o ventrículo esquerdo. Geralmente o S4 marca a contração atrial e a presença dele, ou seja auscultá-lo é sempre sinal de patologia. S3 e S4 têm amplitudes e frequências baixas entre 15-65Hz. A duração desses sons varia de 40 a 60 ms (GOMES; JORGE; AZEVEDO, 2013).

#### 2.1.3.1.2 Murmúrios

Os murmúrios ou sopros cardíacos são distinguíveis dos sons do coração devido à sua duração mais longa. Eles podem ou não ser patológicos. A maioria dos sopros é causada pelo fluxo sanguíneo turbulento que resultados, por exemplo, do estreitamento (estenose) ou vazamento (regurgitação) das válvulas cardíacas ou devido a passagens sanguíneas anormais no coração. De acordo com a situação fisiológica que leva ao sopro, diferentes sons são gerados. Os sopros podem ser sistólicos, diastólicos ou contínuos de acordo à sua localização temporal dentro do ciclo cardíaco. A variação de intensidade do sopro também é importante e os sopros são caracterizados como crescendo, decrescendo, crescendo-decrescendo. Abaixo segue um exemplo de alguns dos murmúrios mais comuns como mostra na Figura 2:

Figura 2 – Exemplos de sopros diastólicos (primeira e terceira divisões), sistólicos (segunda divisão) e contínuos e patologias correspondentes



Fonte: (GHOSH, 2010)

Os murmúrios costumam ter frequências mais altas do que o restante dos componentes de um PCG (200-700Hz) e sua amplitude pode variar de acordo com a gravidade da condição, sendo, por vezes, maior que o primeiro e segundo sons do coração. Além do acima mencionado, outros sons podem ser ouvidos, como a ejeção sistólica precoce ou abertura mitral, que são sons patológicos. Também tem sons não originados no coração, como sons respiratórios, também podem estar presentes no diagnóstico e não podem ser confundidos com patologias (GOMES; JORGE; AZEVEDO, 2013).

## 2.2 Técnicas computacionais

### 2.2.1 Transformada Wavelet

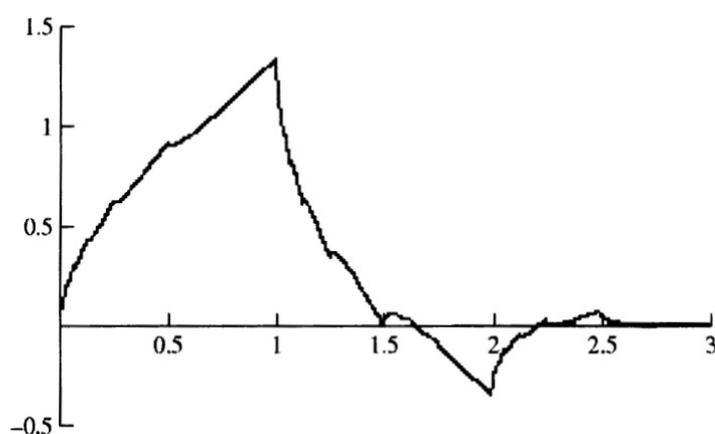
Wavelets são funções capazes de decompor e descrever outras funções originalmente descritas em função de tempo (NIEVERGELT, 2013). Tem frequência variável, duração limitada e valor médio zero.

Em análises com wavelets pode-se extrair informações de domínio da frequência e informações de função de domínio de tempo, diferenciando-se portanto de análise de Fourier. Funções wavelets tem energia contida numa região limitada, ao contrario da análise de Fourier que é periódica e infinita.

Transformadas wavelet decompõe uma função definida no domínio do tempo em outra função, definida no domínio do tempo e frequência. Matematicamente, transformada wavelet pode ser interpretada como convolução do sinal com uma função wavelet (ADDISON, 2017). As transformadas wavelet podem ser contínuas ou discretas.

Dentre a variedade de tipo de wavelets, a do tipo Daubechies é contínua, tendo como consequência dessa característica uma aproximação mais precisa em análise de sinais contínuos utilizando menos wavelets, se comparada com wavelets Haar, por exemplo (NIEVERGELT, 2013).

Figura 3 – Bloco básico de construção de Ingrid Daubechies



Fonte: (NIEVERGELT, 1999)

Para tal wavelet contínua passíveis de transformação dentro de um intervalo com comprimento finito, Ingrid Daubechies introduziu uma função de "bloco básico de construção", mostrada na Figura 3. Para permitir que as wavelets correspondentes analisem partes limitadas de sinais, o bloco de construção é igual a zero fora do intervalo de 0 a 3 (NIEVERGELT, 2013).

Daubechies provou que a função de bloco de construção não admite nenhuma fórmula algébrica em termos de funções matemáticas elementares, no entanto, Daubechies também demonstrou que satisfaz várias relações algébricas que se provam tão úteis quanto fórmulas para fins de cálculos (NIEVERGELT, 2013). A wavelet será usada na etapa de filtragem de sinal.

### 2.2.2 Descritores Sonoros

Utilizado em trabalhos como Kurniawan, Khalil e Malik (2015), Logan et al. (2000) e Hermansky (1990), o MFCC já foi explorado, tendo a eficiência já comprovada na descrição de padrões de fala por ser capaz de expressar a amplitude espectral de forma compacta e fiel. Já o *mel spectrogram* consiste em um artefato obtido durante uma das etapas da obtenção do MFCC (RABINER; JUANG, 1993). O MFCC é obtido submetendo cada dado de áudio a um janelamento dos dados, converte para o domínio a ordem da transformada rápida de Fourier. É então encontrada a magnitude e convertido os dados da transformada em saídas do banco de filtros. Por fim é calculada a base de log 10, e a transformação de cosseno para reduzir a dimensionalidade (SLANEY, 1998).

As características *chroma* do cromatismo são características de áudio obtidas através da projeção de todo o espectro em 12 bins referentes aos semitons musicais, sendo assim uma ferramenta poderosa na descrição de faixas de áudio (ELLIS, 2007). Este nos diz a intensidade de cada um dos semitons em cada espaço de tempo. Após calcular um cromatismo do som em ré o dado é dividido em janelas separadas em intervalos definidos. Então a oitava musical é dividida em passos e pesada com frequência central (em Hz) e função gaussiana (em oitavas) (ELLIS, 2007).

*Spectral contrast*, melhor abordado em Jiang et al. (2002), descreve as frequências obtidas a partir de uma análise baseada em oitavas musicais da frequência do espectrograma dos sinais de áudio. O *Spectral contrast* baseado em oitavas considera o pico espectral, vale espectral e sua diferença em cada sub-banda. A transformada rápida de Fourier é realizada pela primeira vez nas amostras digitais para obter o espectro. Em seguida, o domínio da frequência é dividido em sub-bandas por vários filtros de escala de oitava. A força dos picos espectrais, vales e suas diferenças são estimados em cada sub-mão. Depois de ser traduzido para o domínio Log, o recurso *Spectral contrast* bruto é mapeado para um espaço ortogonal e elimina a relatividade entre as diferentes dimensões pela transformada de Karhunen-Loeve (JIANG et al., 2002).

Proposto por Harte, Sandler e Gasser (2006), o método *tonnetz* detecta variações harmônicas em sinais de áudio, através de análise temporal de alterações de tom em um momento  $n$  a partir da distância euclidiana entre *frames*  $n - 1$  e  $n + 1$  de forma que um pico na função de detecção demarca uma transição de uma região harmonicamente estável à outra. Para a análise é calculado uma transformação de 36 caixas por oitava em cinco

oitavas entre 110Hz e 3520Hz a partir de um sinal de áudio. Um cromatismo sintonizado de 12 bins é então calculado a partir dos espectros dando um vetor chroma de 12 dimensões para cada quadro (HARTE; SANDLER; GASSER, 2006).

A taxa de cruzamento zero é a taxa à qual o valor de um elemento de uma serie temporal de áudio muda de positivo para negativo ao longo do tempo (CHEN, 1988). Esse recurso é utilizado no reconhecimento de fala e na recuperação de informações musicais (GOUYON; PACHET; DELERUE, 2000).

### 2.2.3 Classificação

O Classificador utilizado foi a Máquina de Vetores de Suporte (Support Vector Machine – SVM) (VAPNIK, 1999), já amplamente conhecida e utilizada dentro das mais diversas áreas da inteligência computacional. Máquinas de Suporte de Vetores são algoritmos para classificação eficientes por minimizar superajustes e por suportar muitos atributos (AMARAL, 2016).

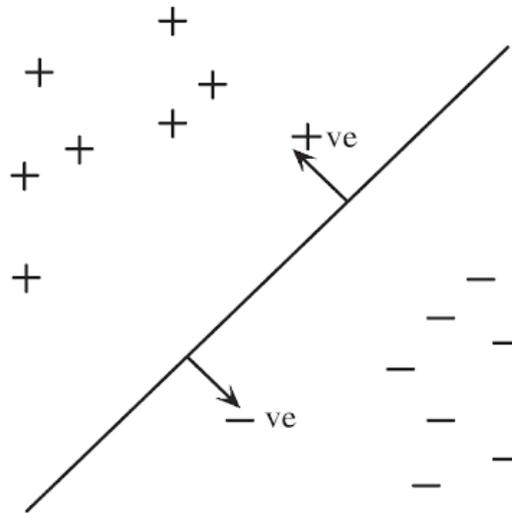
Um SVM é uma máquina de aprendizado abstrata que aprende a partir de um conjunto de dados de treinamento e tenta generalizar e fazer previsões corretas sobre novos dados (CAMPBELL; YING, 2011).

Os dados de treinamento podem ser vistos como pontos rotulados em um espaço que representamos na Figura 4. A tarefa de aprendizado equivale a encontrar um hiperplano direcionado, isto é, um hiperplano orientado de modo que os pontos de um lado sejam rotulados  $y_i = +1$  e aqueles no lado oposto como  $y_i = -1$ . O hiperplano encontrado por um SVM está muito distante das duas classes de pontos rotulados localizados em cada lado. Os pontos mais próximos em ambos os lados têm maior influência na posição deste hiperplano de separação e são, portanto, chamados de vetores de suporte. O hiperplano de separação é dado como  $\vec{w} \cdot \vec{x} - b = 0$ .  $b$  é a polarização ou deslocamento do hiperplano a partir da origem no espaço de entrada,  $x$  são pontos localizados dentro do hiperplano e o normal ao plano, os pesos  $w$ , determinam suas orientações. (CAMPBELL; YING, 2011)

O classificador SVM separa duas classes através de um hiperplano que maximiza uma margem de separação entre os pontos das diferentes classes. Os pontos mais próximos do hiperplano são os de maior influencia da posição do hiperplano e portanto são chamado de vetores de suporte.

Para conjuntos de dados com padrões não-linearmente separáveis, o classificador SVM mapeia os padrões de entrada em um vetor de características com alta dimensão, para que a separação ocorra no novo espaço. Porém, a tarefa de encontrar uma função de transformação de espaço não é trivial. Para resolver o problema, a transformação ocorre com o auxílio de funções não-lineares que são denominadas de Kernel, que torna possível a construção de um hiperplano de separação ótimo para o espaço de características sem

Figura 4 – Exemplo de Hiperplano do SVM



Fonte: (CAMPBELL; YING, 2011)

considerar explicitamente o espaço (SOARES, 2008).

#### 2.2.4 Métricas de Avaliação a Classificação

É esperado que um classificador automático emita respostas corretas para todos os valores e classes da base de dados, mas na prática esse nem sempre é o caso. Se a instância for positiva e classificada como positiva, é contada como um verdadeiro positivo (VP); se for classificada como negativo, é contado como um falso negativo (FN). Se a instância for negativa e classificada como negativa, ela será contada como um verdadeiro negativo (VN); se for classificada como positivo, é contado como falso positivo (FP) (FAWCETT, 2006).

Essas quatro combinações de resultados costumam ser representadas por uma matriz que recebe o nome de "Matriz de Confusão" (QUILICI-GONZALEZ; ZAMPIROLI, 2015) e estão representadas na Tabela 1.

Tabela 1 – Exemplo de Matriz de Confusão

	<b>Positivo</b>	<b>Negativo</b>
<b>Positivo</b>	Verdadeiro Positivo	Falso Negativo
<b>Negativo</b>	Falso Positivo	Verdadeiro Negativo

As Equações das métricas Acurácia (2.1), Precisão (2.2), Sensibilidade (2.3) e F-Measure (2.4) são descritas, respectivamente:

$$\text{Acurácia} = \frac{VP + VN}{VP + VN + FP + FN} \quad (2.1)$$

$$\text{Precisão} = \frac{VP}{VP + FP} \quad (2.2)$$

$$\text{Sensibilidade} = \frac{VP}{VP + FN} \quad (2.3)$$

$$\text{F-Measure} = \frac{2}{1/\text{precisão} + 1/\text{sensibilidade}} \quad (2.4)$$

### 3 Trabalhos Relacionados

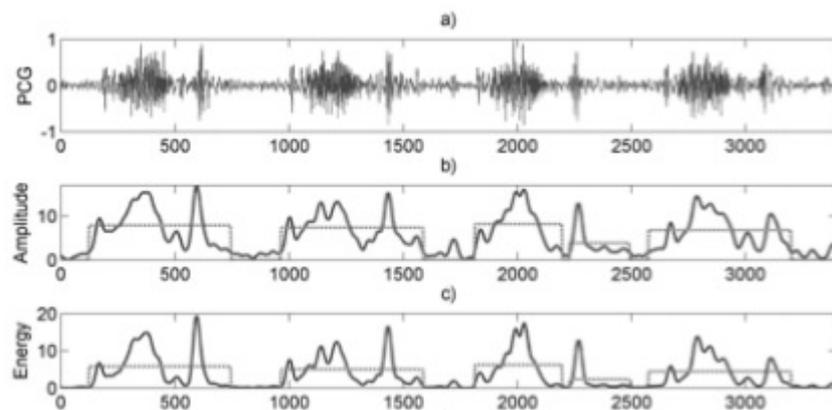
A ausculta assistida por computador tem sido objeto de pesquisa há algum tempo e muitos métodos têm sido aplicados para resolver esta problemática não apenas em termos do algoritmo apropriado e classificação, mas também na maneira como os dados são adquiridos e tratados.

O fonocardiograma é um registro refletindo os sons do coração gerados durante o ciclo cardíaco (MULRONEY SUSAN E MYERS, 2009). O Processamento de sinal de fonocardiograma, ou PCG (do inglês, phonocardiogram), pode ser de modo geral, dividido em duas áreas principais: a primeira focada na detecção de eventos como S1 e S2 para realizar uma segmentação do PCG em ciclos cardíacos, e a segunda com o objetivo de detectar sopros e conseqüentemente patologias cardíacas. Cumpre expor que essas duas áreas de detecção são realizados simultaneamente e estão interligadas a partir do mesmo sinal de base do PCG (DEBBAL; BEREKSI-REGUIG, 2008).

Estas ferramentas de processamento variam, pois há uma grande quantidade de técnicas desde as mais simples até as mais complexas, no entanto, a mais utilizada para ausculta cardíaca é o PCG.

As técnicas mais simples estão compreendidas na análise do domínio do tempo, como as técnicas estatísticas. Técnicas de envelopamento<sup>1</sup> também são bastante usados para simplesmente representar a mudança na amplitude em todo o PCG. Envelopamentos mais complicados, como envelopamentos de energia, Shannon ou Teager também são de uso muito comum. A Figura 5 mostra exemplos da amplitude e energia envolvidas:

Figura 5 – Gráficos de PCG e respectivos envelopes instantâneos de amplitude e energia.



Fonte: Adaptado de Martínez-Alajarín

<sup>1</sup> Método de detecção de oscilação cardíaca frequentemente utilizada antes da classificação final do som cardíaco. metodologia onde a segmentação se baseava nas características de sinal no domínio do tempo

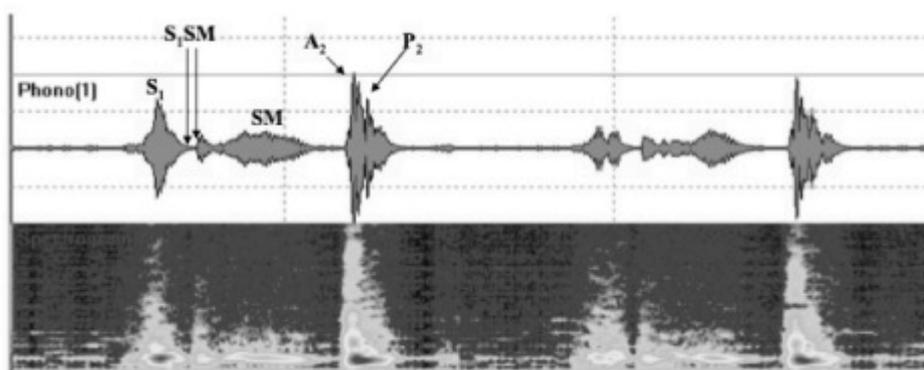
O PCG é uma onda sonora composta por diferentes fontes com diferentes assinaturas de frequência cada um, a análise de domínio de tempo-frequência é uma ferramenta importante para a análise PCG.

Como apontado por Martínez-Alajarín (2008) e o próprio nome faz perceber que, essas técnicas permitem observar a evolução do componentes de frequência do sinal ao longo do tempo. As técnicas utilizadas neste domínio vão desde o envelope de frequência mais simples que representa a frequência principal em um determinado momento para mais métodos complexos que permitem a TFR (Técnicas de Fisioterapia Respiratória) do sinal e, portanto, a observação de seus diferentes componentes de frequência.

Tais técnicas podem ser, por exemplo, a Transformada de Fourier de Curto Termo, Distribuições Wigner-Ville, a Transformada de Gabor e a Transformada Wavelet. A diferença entre estas técnicas em termos de aplicação é a troca entre tempo e frequência. Como essas técnicas originam matrizes de representação de tempo-frequência, ferramentas como valor singular a decomposição são frequentemente usadas para recuperar recursos delas (DELGADO-TREJOS et al., 2009).

A Figura 6 bem como seu espectrograma correspondente obtido pela Transformada de Fourier de curto termo dentro da faixa de frequência de 0 a 1000Hz.

Figura 6 – Gravação do PCG



Fonte: (El-Segaier, 2007)

Um campo um pouco mais reduzido de processamento de PCG é a análise perceptual. Similarmente a muitas outras ferramentas de processamento de PCG, a análise perceptual vinda de técnicas aplicadas em reconhecimento de voz. Ao contrário do processamento digital que trata a frequência em uma dimensão linear, o ouvido humano trata frequência em uma escala logarítmica, a escala Mel. A análise perceptiva é, então, qualquer técnica dependente da conversão do sinal ou do espectro de sinal para a escala Mel. Uma das ferramentas mais usadas neste domínio são *Mel-Frequency Cepstral Coefficients* (MFCC) que permite estudar a forma do espectro de um sinal na escala Mel (DELGADO-TREJOS et al., 2009).

A natureza não-linear e caótica da acústica cardíaca tem sido objeto de muita análise. Sabe-se que os sons S3 e S4 quando são percebidos pelo PCG são considerados sopros, e por tanto dentro a natureza não linear desses sons verifica-se que o S3 e o S4 são os componentes que dão ênfase a essa característica mais caótica. Diante disso verifica-se que os sons cardíacos necessitam de uma ferramenta refinada especialmente em algoritmos.

Os espaços existentes entre os sons também são utilizados como ferramenta importante que dão acesso a análise dos sons não lineares. Assim, a reconstrução do espaço são importante pois proporcionam uma ampla gama de ferramentas de análise não-lineares. Essas ferramentas medem características da trajetória do espaço entre os sons e tal situação enfatiza o caráter não-linear ou caótico do sinal. De maneira generalizada a análise de quantificação de recorrência de espaço entre um som cardíaco e outro, também tem sido utilizada para caracterizar o estado em que se encontra o coração (ERGEN; TATAR; GULCUR, 2012).

Um esforço recente também está sendo feito na separação dos diferentes componentes do PCG. As metodologias de separação de fontes cegas incluem, por exemplo, a análise de componentes principais e a decomposição de valor singular. No entanto, para que tal abordagem seja possível, há a necessidade de um número de sinais igual ou maior ao número de fontes. No entanto, os resultados são bastante promissor como mostrado na Figura 7.

O método de aquisição do sinal PCG é tão importante quanto o algoritmo de processamento implementado e tem implicações para todo o algoritmo e aplicação dos resultados.

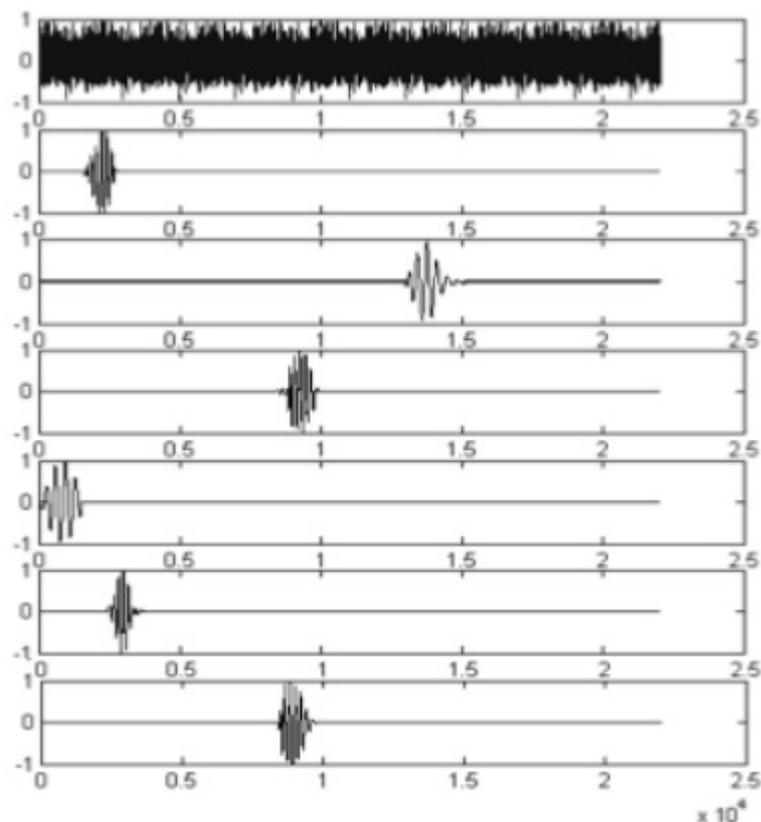
Uma das abordagens utilizadas é a aquisição do sinal PCG simultaneamente com outro bio-sinal. O mais comum é o eletrocardiograma (ECG), mas outros sinais, como a frequência respiratória ou o ecocardiograma foram relatados. A tendência de aquisição simultânea de PCG e ECG é comum devido à sincronização entre os dois sinais. Como mostrado na Figura 8 S1 corresponde ao tempo para o complexo QRS <sup>2</sup> no ECG e S2 segue a pausa sistólica no ciclo cardíaco normal.

A abordagem alternativa é a única aquisição do PCG sem outro bio-sinal. Esta abordagem é mais consensual do que qualquer resultado obtido através deste método extrapolado para uma prática clínica, onde o médico iria adquirir o PCG com um estetoscópio digital e obter os resultados em tempo real.

A detecção do som do coração é muito mais difícil e novos algoritmos são necessários. Resultados promissores foram alcançados com bancos de dados PCG adquiridos em ambientes controlados, como será mostrado no capítulo 4 que trata da metodologia

<sup>2</sup> A sigla refere-se a despolarização ventricular sendo que o primeiro componente é relativo ao aurículo direito e o segundo ao aurículo esquerdo. Representam batidas por hora, se um indivíduo normal tem 80 batidas por minuto o QRS será calculado assim:  $60 \times 80$ , que resulta em 4800 QRS em uma hora.

Figura 7 – Os componentes separados de um sinal normal de PCG. De cima para baixo: ruído de fundo, componente mitral de S1, S3, componente aórtico de S2, S4, componente tricúspide de S1 e componente pulmonar de S2.



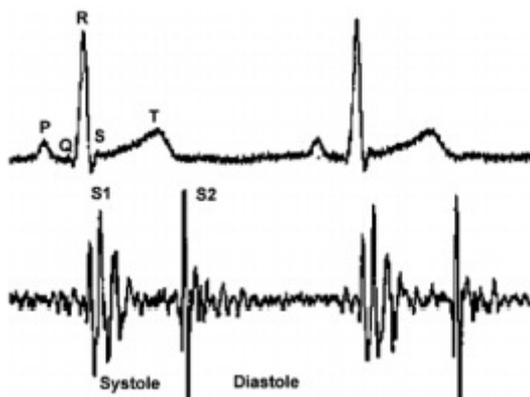
Fonte: (NIGAM; PRIEMER, 2005)

utilizada no presente estudo. O desafio é então aplicar algoritmos a bancos de dados compostos de sinais de PCG adquiridos em ambiente clínico real e em procedimentos semelhantes aos métodos de auscultação utilizados pelos médicos.

Há uma grande variedade de algoritmos de segmentação disponíveis na literatura, e pela variedade existente, a descrição dos métodos muitas vezes são impraticáveis, como consequência, nem todos os métodos são utilizados. Portanto, poucos foram os métodos escolhidos para serem descritos, como segue.

El-Segaier et al. (2005) desenvolveu um método baseado no controle de ECG. Este método usa o método simultâneo aquisição de sinais de PCG e ECG. Usando um algoritmo de detecção baseado em envelopamento, as ondas R do complexo QRS do ECG foram detectadas e a distância R e ondas T são calculadas. A transformada de Fourier de curto termo foi então usada para obter o espectro do PCG. Usando as relações temporais entre o PCG e o ECG, foram definidos intervalos de busca para S1 e S2 e os máximo no espectro em cada um desses intervalos foram definidos como S1 ou S2. Uma ferramenta adicional foi também desenvolvido para determinar se a obtenção máxima formava um pico bem

Figura 8 – ECG sincronizado (superior) e PCG (inferior) mostrando o tempo do complexo QRS-S1 e T-onda-S2 relações



definido no domínio do tempo.

Martinez-Alajarin (2005) desenvolveram um método de segmentação dependente principalmente da análise no domínio do tempo. Usando os envelopes de amplitude, energia e frequência, a frequência cardíaca é encontrada através do ACF, uma função simples especializada em encontrar os elementos periódicos de um sinal. O envelope de amplitude e uma série de regras definidas empiricamente é usado para encontrar os eventos de interesse, o coração soa.

Ainda acerca dos algoritmos, vale destacar o estudo de Naseri e Homaeinezhad (2013) em que o estudioso, desenvolveu um método baseado tanto no domínio do tempo quanto na análise deste. Uma função específica foi projetada para ser sensível a altas amplitudes e frequências específicas dos principais sons do coração. Isso foi feito usando a Transformada Rápida de Fourier. Os picos da função de envelopamento obtidos foram então considerados como candidatos a sons do coração. Deles forma e duração foram também avaliadas e, se validadas de acordo com estes parâmetros, os eventos poderia ser classificado em S1 ou S2. Este processo foi realizado iterativamente ao longo do sinal PCG.

Gill, Gavrieli e Intrator (2005) desenvolveu um algoritmo usando mapas probabilísticos auto-organizáveis. Um filtro é usado para obter um envelope fino. Este método lida com picos divididos ou serrilhados usando um suavidade escalável. Os picos do envelopamento são então usados em um HMM usando como observações amplitude do pico, a distância temporal entre os picos adjacentes, suas amplitudes e a segunda derivada do pico.

Oskiper e Watrous (2002) projetaram uma rede neural com atraso de tempo para realizar a segmentação tarefa. Utilizaram-se 40 escalas de wavelets que englobaram as frequências 10-299Hz para obter uma matriz TFR de PCG nas frequências dos principais sons do coração. Um ECG adquirido simultaneamente é usado como ponto fiducial e duas redes neurais de retardo de tempo diferentes são treinadas.

A primeira variação, e o ponto de partida de todo o processo, é a escolha do banco

de dados. Nesse caso, diferentemente da segmentação, o banco de dados relaciona-se à aquisição simultânea de outros sinais na intenção de compará-los. A escolha do banco de dados é relacionada à maneira como a aquisição desses dados será realizada, isto é, as condições em que o estudo dos sons será conduzidos. Assim, ao escolher o bando de dados deve ser levado em consideração os sujeitos do estudo, o ambiente e precauções tomadas e a quantidade de informações “verdadeiras” disponíveis sobre cada sinal. Cabe enfatizar, que mesmo que o banco de dados utilize outros sinais no algoritmo de segmentação estes apenas servem para o conhecimento do autor e nem todos serão utilizados na detecção de murmúrios.

Ainda acerca da escolha do banco de dados, [Dokur e Ölmez \(2008\)](#) apontam que o uso exclusivo de PCGs simulados por computador tiveram claras alterações quanto a aplicação do resultado final. Essas alterações se devem principalmente pelo tamanho do banco de dados utilizado, uma vez que a quantidade de informação é muito maior do que a esperada. Assim, quanto maior for o banco de dados maior também será a variação de sinais e, portanto, seus resultados serão muito mais confiáveis.

Um ponto a considerar no processo de classificação é que apesar da pouca importância dada no momento da classificação, este é um processo de extrema relevância, principalmente por que a classificação incorreta pode acarretar na alteração dos sinais obtidos.

O número e o tipo de classes também são variáveis. Existem dois métodos principais de divisão. Um visa identificar a patologia exata que está presente e, portanto, apresenta uma classe para cada uma das patologias encontradas no banco de dados. A segunda visa identificar se existe ou não um sopro presente independente da patologia que originou isso. A primeira abordagem requer obviamente um conhecimento muito mais profundo dados clínicos do paciente ([CARVALHO et al., 2011](#)).

A seguir, o presente estudo apresenta os algoritmos de classificação que darão melhor compreensão de como essas classificação deverão ser trabalhadas.

### 3.1 Algoritmos de Classificação Cardíacos

Um dos métodos que atende a todos os padrões é a fonocardiografia (PCG), definida como o monitoramento do sistema circulatório humano, registrando a atividade biomecânica do coração. Apesar de sua simplicidade e facilidade de implementação, raramente é usado para o diagnóstico cardíaco, principalmente devido à falta de soluções razoáveis, que poderiam permitir a interpretação inequívoca dos resultados. Por outro lado, esta técnica pertence a um grupo de métodos cujo desenvolvimento é particularmente necessário em sistemas de auto-análise (como estetoscópios inteligentes).

A pesquisa atual neste campo é focada no desenvolvimento de algoritmos adequados, que no futuro podem levar ao desenvolvimento de um estetoscópio inteligente. Devido à natureza dos sinais de PCG e ao ruído indesejado durante o exame, é importante dividir o processo de diagnóstico em duas etapas. O primeiro é o processamento de sinais originais destinados a extrair características, o que ajudaria a distinguir todos os tipos de sinais e o segundo associado ao processo de classificação de sinais, onde a maioria dos trabalhos foram focados na aplicação de técnicas baseadas em redes neurais artificiais e máquinas de vetores de suporte (BENTLEY et al., 2011).

Uma decomposição de wavelets multinível com uma perceptron multicamadas treinada por um algoritmo de propagação reversa alcançou 94,42% de precisão na identificação de quatro estados do coração (KUMAR et al., 2008). Outros trabalhos incluem o uso de busca multivariada para modelar sopros e classificá-los com uma rede perceptron de três camadas com 92,5% de precisão (distinguindo estados cardíacos normais de anormais). Martinez-Alajarin (2005) usou uma combinação de detecção de características cardíacas (atividade, complexidade, mobilidade e picos espectrais) com RNA, proporcionando uma taxa de 98% de identificação, porém capaz de distinguir apenas três deles.

Outro grupo de métodos emprega o SVM como o classificador principal. Uma abordagem para identificação de sons cardíacos apresentada por Castro (et al 2013) garantiu 95% de precisão usando o som transformado wavelet para extrair o envelope de sinais PCG. No entanto, os autores conseguiram distinguir apenas estados cardíacos normais e anormais. Os mesmos resultados foram alcançados por Ghosh (2010), que desenvolveram um sistema para uso doméstico, no entanto, este sistema foi comprovado apenas por um estudo de caso.

Neste caso, o sistema assegurou uma precisão de 96% na classificação normal e 93,1% de dois estados cardíacos anormais. Mais tarde, a Avci melhorou este sistema e desenvolveu o Gene Support Vector Machines, que obteve 95% de precisão (DELGADO-TREJOS et al., 2009).

Outros trabalhos notáveis introduziram soluções como: um modelo analítico, baseado em um único DOF para extrair formas de onda características de sons cardíacos e um método de agrupamento C-means fuzzy para sua classificação (GHOSH, 2010); análise de tempo-frequência em conjunto com entropia de Rényi (NASERI; HOMAEINEZHAD, 2013) ou coeficientes cepstrais de frequência Mel com modelos ocultos de Markov (TAVEL, 2006).

Diante do exposto, verifica-se que, os principais problemas relacionados com o desenvolvimento de técnicas relevantes eram em sua grande maioria a questão da variedade de sons cardíacos patológicos distinguíveis e o caráter não estacionário dos sinais de PCG. Assim, como a qualidade dos sinais da fonocardiografia aumentou consideravelmente não é mais um problema a dificuldade na percepção dos sons cardíacos, pois o desenvolvimento

de algoritmos apropriados, capazes de distinguir um grande número de doenças cardíacas com base em sons cardíacos relevantes, tornou-se uma questão essencial que alcançou sistemas capazes de perceber com antecedência sons cardíacos com problemas.

Tendo em mente essas questões, é levantada uma questão sobre como aumentar a variedade de sons cardíacos distintos e melhorar o desempenho de tais sistemas em termos de redução de sua complexidade computacional sem comprometer a precisão. Portanto, faz-se essencial a utilização de um novo método para identificação de sons cardíacos como se propõe apresentar no presente estudo monográfico.

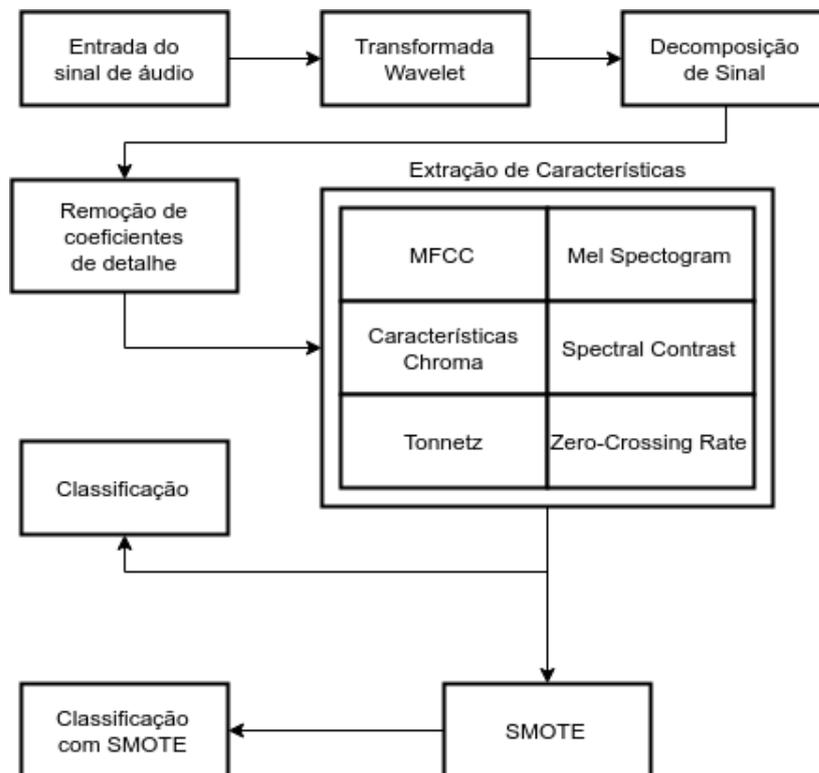
Sobre esse aspecto cabe apresentar com brevidade um algoritmo que representa de maneira sucinta como deve ocorrer o tratamento desses sons.

Diante do diagrama acima exposto, é perceptível que o sinal PCG deve ser processado adequadamente e todos os recursos característicos extraídos para fornecer informações suficientes para os classificadores. A maioria dos trabalhos relacionados a esse assunto [Tavel \(2006\)](#) e [El-Segaier et al. \(2005\)](#) empregaram vários tipos de análise de wavelets, que é uma técnica computacionalmente complexa. Neste trabalho, é mostrado que, usando a análise espectral, é possível fornecer informações suficientes para o classificador e manter pouca complexidade computacional.

## 4 Materiais e Método

Este capítulo fornece uma descrição dos dados usados e os método usado para tratá-los. Assim o presente capítulo foi dividido em: informações sobre a base de dados utilizada, as rotinas de processamento do sinal de áudio e classificação. O diagrama da Figura 9 ilustra as etapas do método proposto.

Figura 9 – Etapas do método proposto



Fonte: Acervo do autor

### 4.1 Descrição da Base de Dados

Os arquivos de áudio usados no estudo provem de *The PASCAL Classifying Heart Sounds Challenge 2011* (BENTLEY et al., 2011). Essa base de dados é dividida em duas base de dados menores denominadas de A e B. A base de dados A contém 176 arquivos de áudio de exames cardíacos no formato WAV, sendo 31 rotulados como normal, 34 como sopro, 19 como sons cardíacos acessórios, 40 como artefato, onde não há sons de coração discerníveis e 52 sem rotulação. Foram adquiridos utilizando o aplicativo para iPhone iStethoscope Pro. A base de dados B contém 656 arquivos no formato WAV, sendo 320 arquivos rotulados como normal, 95 como sopro cardíaco, 46 como extrassístole

ventricular e 195 sem rótulo. Foram adquiridos através de teste clínico em hospitais usando o estetoscópio digital DigiScope. Em ambas as base de dados os arquivos tem duração entre 1 e 30 segundos.

As imagens da Figura 10 exemplificam os sinais utilizados da base de dados. A figura 10a representa um sinal de áudio de batimento de um paciente saudável. A figura 10b representa o sinal de áudio de um batimento de um paciente com murmúrios ou sopros cardíacos. A figura 10c representa um sinal de áudio de um batimento diagnosticado com extrassístole Sopros cardíacos tem sua duração mais longa que o som de um batimento normal e podem ou não ser patológicos. A maioria dos sopros é causada pelo fluxo sanguíneo turbulento resultados do estreitamento ou vazamento das válvulas cardíacas ou devido a passagens sanguíneas anormais no coração.

Na filtragem da base de dados cada arquivo de áudio foi processado como um sinal que foi decomposto utilizando a Transformada Wavelet Discreta Multinível de uma dimensão. O processo de decomposição gera dois tipos de coeficientes: coeficientes de aproximação e o coeficiente de detalhe. Neste trabalho o coeficiente de detalhe foi removido e apenas foi utilizado os coeficientes de detalhe.

A wavelet utilizada foi da família Daubechies de ordem 6 de 4 níveis, exemplificada na Figura 11. Durante a filtragem todos os coeficientes de detalhe foram removidos e então o sinal foi reconstruído.

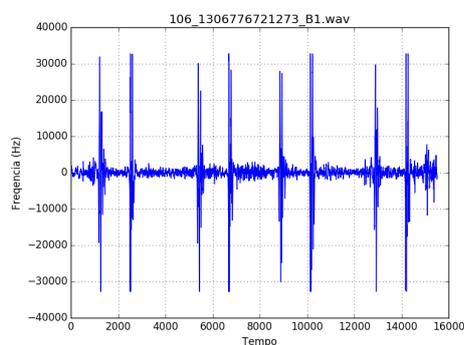
Foram extraídas 194 características acústicas obtidas a partir de 6 descritores de artefatos em áudio, já descritos no Capítulo 2.2 - 40 características *mel-frequency ceptral coefficient* (MFCC), um vetor de 12 características *chroma, mel spectrogram* de 128 características, 7 características *spectral contrast, tonnetz* com 6 características e um *zero-crossing rate* de série temporal de áudio.

Então é feita a rotulação das instâncias de acordo com o sinal fonte da base de dados. A classificação foi realizada utilizando a máquina de vetores de suporte (SVM), já amplamente conhecida e utilizada dentro das mais diversas áreas da inteligência computacional (SOARES, 2008) (VAPNIK, 1999).

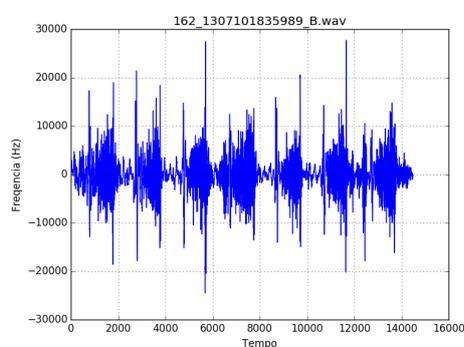
## 4.2 Rotinas de Extração de Características

Esse passo consiste em decompor o sinal original usando a Transformada Wavelet como mostra a Figura 9 e filtragem para remoção de ruído, baseado nas metodologias propostas em (GOMES; PEREIRA, 2012) e (DENG; BENTLEY, 2012)

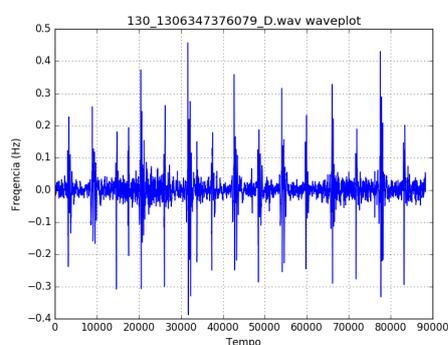
A etapa de extração de características foi realizada com a utilização da biblioteca python *librosa* (MCFEE et al., 2018), que é uma biblioteca específica desenvolvida para a análise de faixas de áudio que permite a extração de características sobre estes sinais.

Figura 10 – Exemplos de áudios originais de *The PASCAL Classifying Heart Sounds Challenge 2011*

(a) Áudio de batimento normal



(b) Áudio de batimento com sopro



(c) Áudio de batimento extrassístole

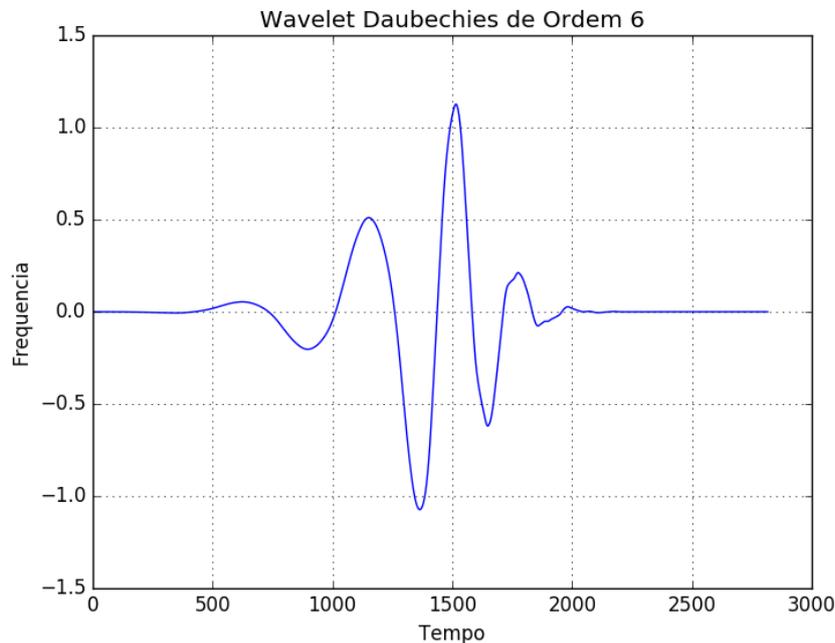
Fonte: Acervo do autor

As etapas de normalização e classificação foram efetuadas utilizando o software WEKA (FRANK; WITTEN, 2016).

### 4.3 Normalização e Re-amostragem de Instâncias

Após o processo de extração de características e rotulação de acordo com a base de dados, todas as características foram normalizadas para valores entre 0 e 1.

Figura 11 – Wavelet Daubechies de Ordem 6



Fonte: Acervo do autor

Após a normalização também foi realizada a re-amostragem utilizando a técnica de Sobreposição de Minoria Sintética (*Synthetic Minority Oversampling TEchnique* – SMOTE) para balancear a base de dados e obter melhor desempenho do classificador (AL., 2002). Nessa técnica a classe minoritária é sobre-amostrada criando exemplos "sintéticos" em vez de super-amostragem com substituição. A classe minoritária é sobre-amostrada tomando cada amostra de classe minoritária e introduzindo exemplos sintéticos ao longo dos segmentos de linha unindo qualquer ou todos os vizinhos mais próximos da classe minoritária (AL., 2002). Neste projeto a quantidade de instancias para base A aumentou em 19 e, para base B aumentou 204 instancias.

Amostras sintéticas são geradas da seguinte maneira: É feita a diferença entre o vetor de característica em consideração e seu vizinho mais próximo, multiplica-se essa diferença por um número aleatório entre 0 e 1 e adicione-o ao vetor de características em consideração (AL., 2002). A Tabela 2 exemplifica a geração de exemplos sintéticos.

Tabela 2 – Exemplo de geração de exemplos sintéticos.

---

Considere uma amostra (6,4) e seja (4,3) o seu vizinho mais próximo.  
(6,4) é a amostra para a qual os k vizinhos mais próximos estão sendo identificados.  
(4,3) é um dos seus vizinhos com k mais próximos.  
Sendo assim:  
 $f1_1 = 6$   $f2_1 = 4$   $f2_1 - f1_1 = -2$   
 $f1_2 = 4$   $f2_2 = 3$   $f2_2 - f1_2 = -1$   
As novas amostras serão geradas como  
 $(f1', f2') = (6,4) + \text{rand}(0-1) * (-2, -1)$   
O rand (0-1) gera um número aleatório entre 0 e 1.

---

## 4.4 Classificação

O Classificador utilizado foi a Máquina de vetores de suporte (Support Vector Machine - SVM) (VAPNIK, 1999), já amplamente conhecida e utilizada dentro das mais diversas áreas da inteligência computacional. Máquinas de Suporte de Vetores são algoritmos para classificação eficientes por minimizar superajustes e por suportar muitos atributos (AMARAL, 2016).

Os métodos do kernel são uma classe de algoritmos para análise de padrões, como por exemplo, a máquina de vetores de suporte. A tarefa geral da análise de padrões é encontrar e estudar tipos gerais de relações em conjuntos de dados. O SVM foi usado como método de kernel de um algoritmo de pesquisa em grade. A pesquisa de grade executa a pesquisa de pares de parâmetros para o classificador e escolhe o melhor par encontrado para a previsão real. Uma vez que o melhor ponto da grade é obtido é executada uma validação cruzada com os pares de parâmetros adjacentes. Se um par melhor for encontrado, então isso funcionará como um novo centro e outra validação será realizada. Esse processo é repetido até que nenhum par melhor seja encontrado ou o melhor par esteja na borda da grade (PFAHRINGER; HOLMES, 2014).

## 5 Resultados e Discussão

Neste capítulo são apresentados os resultados obtidos pelo método apresentado no Capítulo 4. Os resultados estão divididos em duas seções. A primeira exibe os resultados obtidos com a filtragem dos sinais utilizando a Transformada Wavelet e a normalização das características. Por fim são apresentados os resultados de testes feitos na re-amostragem de instâncias utilizando SMOTE, explicada no Capítulo 4.

Todos os resultados exibidos a seguir foram obtidos com validação cruzada com 10 partições. Com as características extraídas a partir de 6 descritores de artefatos em áudio utilizados, descritos no Capítulo 2.2, – *mel-frequency ceptral coefficient* (MFCC), *mel spectrogram*, *chroma features*, *spectral contrast*, *tonnetz* e *zero-crossing rate* – após a filtragem do sinal obteve-se a taxa de acurácia de 66,93% para a Base A, com 83 instâncias classificadas corretamente e 65,7% com 205 instâncias classificadas corretamente para a Base B, apresentados nas Tabelas 3, 4, 5, 6:

Tabela 3 – Matriz de Confusão dos resultados da classificação após filtragem da Base A

Classificação	Normal	Som extra do coração	Murmúrio	Artefato
Normal	16	6	4	5
Som extra do coração	11	8	0	0
Murmúrio	8	2	23	1
Artefato	3	0	1	36

Tabela 4 – Precisão detalhada por classe após filtragem da base A

Base A	Precisão	Sensibilidade	F-Measure
Normal	0,421	0,516	0,464
Som extra do coração	0,500	0,421	0,457
Murmúrio	0,821	0,676	0,742
Artefato	0,857	0,900	0,878

Tabela 5 – Matriz de Confusão dos resultados da classificação após filtragem da Base B

Classificação	Normal	Extra-sístole	Murmúrio
Normal	189	2	9
Extra-sístole	45	0	1
Murmúrio	50	0	16

Tabela 6 – Precisão detalhada por classe após filtragem da base B

Base B	Precisão	Sensibilidade	F-Measure
Normal	0,665	0,945	0,781
Extra-sístole	0,000	0,000	0,000
Murmúrio	0,615	0,242	0,348

Nesses resultados observa-se que a classe Som extra do coração da Base A obteve menor sensibilidade e foi a classe que apresentou mais casos de falha. Essa classe também é a menos presente, contando com apenas 19 instâncias. Apesar disso a classe Normal obteve a menor precisão. Na Base B a classe Extra-sístole foi classificada completamente de maneira equivocada, tendo zero como precisão e todas suas instâncias classificadas como Normal na matriz de confusão. Assim como a classe Som extra do coração da Base A, a classe Extra-sístole é também a com menor representatividade na base de dados, com 46 instâncias.

Por outro lado, na Base A a classe Murmúrio obteve maiores índices de precisão, sensibilidade e *F-measure* dentre as classes que representam um estado preciso do som do coração. A classe Artefato obteve maiores valores de precisão detalhada, mas representa sinais de áudio com muito ruído. Na Base B a classe com maior acerto é a classe Normal, obtendo os maiores valores de precisão, sensibilidade e *F-measure*.

## 5.1 Resultados obtidos com geração de exemplos sintéticos (SMOTE)

Também foi feita a re-amostragem de instâncias utilizando a técnica SMOTE, explicada no Capítulo 4. Essa re-amostragem alterou a quantidade de instâncias associadas a cada classe com a finalidade de balancear as classes na base de dados. É descrito a seguir nas Tabelas 7 e 10. Os testes realizados após esse processo obtiveram acurácia de 78,32% (112 instancias classificadas corretamente) para base A e 84,1% (434 instancias classificadas corretamente) para base B, como mostrado nas Tabelas 8, 9, 11 e 12:

Tabela 7 – Re-amostragem das instâncias da Base A.

Re-amostragem	Antes	Depois
Normal	31 Instâncias	31 Instâncias
Som extra do coração	19 Instâncias	38 Instâncias
Murmúrio	34 Instâncias	34 Instâncias
Artefato	40 Instâncias	40 Instâncias

Tabela 8 – Matriz de Confusão dos resultados da classificação após filtragem e SMOTE na Base A.

<b>Classificação</b>	<b>Normal</b>	<b>Som extra do coração</b>	<b>Murmúrio</b>	<b>Artefato</b>
<b>Normal</b>	17	9	3	2
<b>Som extra do coração</b>	0	38	0	0
<b>Murmúrio</b>	8	4	21	1
<b>Artefato</b>	2	1	1	36

Tabela 9 – Precisão detalhada por classe após filtragem, normalização e SMOTE da base A

<b>Base A</b>	<b>Precisão</b>	<b>Sensibilidade</b>	<b>F-Measure</b>
<b>Normal</b>	0,630	0,548	0,586
<b>Som extra do coração</b>	0,731	1,000	0,844
<b>Murmúrio</b>	0,840	0,618	0,712
<b>Artefato</b>	0,923	0,900	0,911

Tabela 10 – Re-amostragem das instancias da Base B

<b>Re-amostragem</b>	<b>Antes</b>	<b>Depois</b>
<b>Normal</b>	200 Instâncias	200 Instâncias
<b>Extra-sístole</b>	46 Instâncias	184 Instâncias
<b>Murmúrio</b>	66 Instâncias	132 Instâncias

Tabela 11 – Matriz de Confusão dos resultados da classificação após filtragem e SMOTE da Base B

<b>Classificação</b>	<b>Normal</b>	<b>Extra-sístole</b>	<b>Murmúrio</b>
<b>Normal</b>	175	12	13
<b>Extra-sístole</b>	15	169	0
<b>Murmúrio</b>	40	2	90

Tabela 12 – Precisão detalhada por classe após filtragem, normalização e SMOTE da base B

<b>Base B</b>	<b>Precisão</b>	<b>Sensibilidade</b>	<b>F-Measure</b>
<b>Normal</b>	0,761	0,875	0,814
<b>Extra-sístole</b>	0,923	0,918	0,921
<b>Murmúrio</b>	0,874	0,682	0,766

Após esse procedimento a classe Normal da Base A obteve menor índice de precisão, sensibilidade e *F-measure*, mas ainda melhores que os resultados apenas com a filtragem. A classe Som extra do coração também obteve melhores índices. Na base B, a classe Normal obteve o menor valor de precisão e a classe Murmúrio teve os menores valores de sensibilidade e *F-measure*, sendo essa a classe com mais erros na matriz de confusão.

Da Base A a classe com maior precisão foi a classe Murmúrio. Repetindo o ocorrido dos resultados obtidos apenas com a filtragem a classe Artefato apresentou o maior índice de precisão. Na Base B a classe com melhores índices de precisão, sensibilidade e *F-measure* foi a classe Extra-sístole. Esse resultado vai de encontro aos resultados anteriores onde essa classe obteve os piores resultados.

## 5.2 Resultados com seleção de características

A seleção de características visa identificar as características relevantes para obter um desempenho de classificação semelhante ou até melhor do que o uso de todas as características. Utilizando a implementação do algoritmo por otimização de enxame de partição (*Partition Swarm Optimization* - PSO) para seleção dos atributos pode-se reduzir o número de características de 194 para 25. Realizando os mesmo testes anteriores obteve-se os resultados a seguir.

Para a base de dados A o algoritmo de seleção de características extraiu 37 atributos do MFCC, todo o vetor chroma, 124 atributos de mel, 5 atributos de contrast e 3 de tonnez. Na base B foram removidos 37 atributos de MFCC, todo o vetor *chroma*, 111 atributos de mel, 5 atributos de *spectral contrast* e de *tonnez*.

Tabela 13 – Resultados após Seleção de Características

<b>Classificações Corretas / Porcentagens</b>	<b>Filtragem e Normalização</b>	<b>Filtragem, Normalização + SMOTE</b>
<b>Base A</b>	72 / 58,06 %	87 / 60,83 %
<b>Base B</b>	210 / 67,30 %	405 / 78,48 %

Dentre os resultados dos testes na base de dados A a classe que mais apresentou erros foi a classe Normal. Todas as classes nessa base de dados tem número semelhante de instâncias. Por mais que essa classe apresente mais erros, esse valor está próximo das demais classes da base. Para classes da base B foi a classe Extra-sístole que mais apresentou erros de classificação, possivelmente por estar em menor quantidade em relação as outras classes.

Os melhores resultados apresentados para ambas as bases foram obtidos com a utilização da técnica de re-amostragem SMOTE, com 78,32% de acurácia para base A

e 84,1% de acurácia para base B. Já os piores resultados para base A foram observados utilizando a seleção de características, com uma acurácia de 58,06%. Para base B a simples abordagem básica foi a que apresentou os piores resultados, com 65,7% de acurácia.

Outras abordagens que poderiam melhorar os resultados obtidos poderiam realizar uma segmentação dos sinais de áudio após o pre-processamento, utilizando algoritmos de detecção dos picos nos fonocardiogramas para identificar batimentos cardíacos presentes nos sinais de áudios.

### 5.3 Comparação de resultados

O artigo de [Gomes e Pereira \(2012\)](#), utilizando a mesma base de dados desse trabalho e as técnicas de classificação *Multilayer Perceptron* (MLP) e árvores de decisão geradas pela classe J48. Os resultados obtidos pelo artigo são comparados com os obtidos nessa monografia e estão nas tabelas a seguir 14 e 15, sob as colunas J48 e MLP. Se observa que o classificador utilizado nessa monografia obteve melhores resultados para precisão na classificação das classes.

Tabela 14 – Comparação de resultados com [Gomes e Pereira \(2012\)](#), de [Deng e Bentley \(2012\)](#) e da monografia na Base A

Dataset A	J48	MLP	Artigo de Deng e Bentley	Classificação com filtragem e SMOTE
Precisão de Normal	25%	35%	45,83%	63%
Precisão de Murmúrio	47%	67%	31,25%	84%
Precisão de Som extra do coração	27%	18%	11,27%	73,1%
Precisão de Artefato	71%	92%	58,33%	92,3%
Sensibilidade de Artefato	63%	69%	43,75%	90%

Tabela 15 – Comparação de resultados com [Gomes e Pereira \(2012\)](#), de [Deng e Bentley \(2012\)](#) e da monografia na Base B

Dataset B	J48	MLP	Artigo de Deng e Bentley	Classificação com filtragem e SMOTE
Precisão de Normal	72%	70%	77,67%	76,1%
Precisão de Murmúrio	32%	30%	36,99%	87,4%
Precisão de Extra-sístole	33%	67%	16,67%	92,3%

O artigo de [Deng e Bentley \(2012\)](#), também utilizando da mesma base de dados, obteve os resultados expressos também nas tabelas 14 e 15, também comprando resultados obtidos pelo estudo monográfico. A classificação feita nesse artigo utilizou métodos próprios, onde os autores segmentavam o sinal de áudio, identificavam a quantidade de batimentos cardíacos e características do período de sístole e diástole, e para a base B, era também identificada o comprimento dos picos de áudio. Na base de dados A, a classificação

---

feita nesse trabalho monográfico apresenta melhores resultados de precisão por classes, entretanto na base de dados B o artigo apresenta melhor classificação para classe Normal.

## 6 Conclusão

Este trabalho apresentou um método automático para classificar sinais de áudios de batimentos cardíacos, obtidos através da base de dados *The PASCAL Classifying Heart Sounds Challenge* 2011 (BENTLEY et al., 2011). O método proposto realiza a decomposição e filtragem utilizando Transformada Wavelet. Em seguida, realiza-se a extração de características utilizando descritores de áudio nos sinais filtrados. Por fim, utilizou-se a Máquina de Vetor de Suporte como classificador.

Foram utilizados os seguinte descritores: *mel-frequency ceptral coefficient* (MFCC), *mel spectrogram*, *chroma features*, *spectral contrast*, *tonnetz* e *zero-crossing rate*. O conjunto de características foi extraído sobre o sinal de áudio de cada paciente. Após esse processo realizou-se a normalização das características extraídas de cada instancia de sinal re-amostragem dos dados.

Por fim, essas características são submetidas ao classificador SVM. Dentre essas abordagens, os melhores resultados, tanto para a base A, quanto para base B foram obtidos após a filtragem, normalização e re-amostragem das instâncias, com acurácia de 78,32 para a base A e 84,1% para a base B.

A metodologia proposta apresentou a ideia de extrair características de um sinal filtrado e realizou a análise e classificação desse novo sinal. Assim, a principal contribuição deste trabalho está em oferecer uma ferramenta para auxiliar opiniões de um médico especialista.

Por fim, como trabalho futuro, sugerimos explorar novas formas de representar as gravações de auscultação, fazendo com que estas venham a ser melhor representadas com maior fidelidade e aumentando a confiabilidade de suas características. Precisa-se, também, investigar uma organização dos dados apresentados aos classificadores de forma que este seja capaz de identificar com maior confiança, dado que uma base desequilibrada influencia negativamente na etapa de aprendizado.

# Referências

- ADDISON, P. *The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance, Second Edition*. CRC Press, 2017. ISBN 9781315355283. Disponível em: <<https://books.google.com.br/books?id=VrTZDQAAQBAJ>>. Citado na página 24.
- AL., N. V. C. et. Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, v. 16, p. 321–357, 2002. Citado na página 40.
- AMARAL, F. *Introdução o à Ciencia de Dados: mineração de dados e big data*. ALTA BOOKS, 2016. ISBN 9788576089346. Disponível em: <<https://books.google.com.br/books?id=hAIVDQAAQBAJ>>. Citado 2 vezes nas páginas 26 e 41.
- AMIRI, A. M.; ARMANO, G. Heart sound analysis for diagnosis of heart diseases in newborns. *APCBEE procedia*, Elsevier, v. 7, p. 109–116, 2013. Citado 2 vezes nas páginas 16 e 17.
- BENTLEY, P. et al. *The PASCAL Classifying Heart Sounds Challenge 2011 (CHSC2011) Results*. 2011. [Http://www.peterjbentley.com/heartchallenge/index.html](http://www.peterjbentley.com/heartchallenge/index.html). Citado 5 vezes nas páginas 21, 22, 35, 37 e 48.
- BRASIL, P. *Cerca de 300 mil brasileiros morrem de doenças cardiovasculares por ano*. 2018. Disponível em: <<http://www.brasil.gov.br/saude/2017/09/cerca-de-300-mil-brasileiros-morrem-de-doencas-cardiovasculares-por-ano>>. Citado na página 13.
- CAMPBELL, C.; YING, Y. *Learning with Support Vector Machines*. Morgan & Claypool, 2011. (Synthesis Lectures on Artificial Intelligence). ISBN 9781608456161. Disponível em: <<https://books.google.com.br/books?id=uhqmlu0lgf8C>>. Citado 2 vezes nas páginas 26 e 27.
- CARVALHO, P. de et al. A framework for acoustic cardiac signal analysis. In: . SciTePress, 2011. p. 151–160. Disponível em: <[https://www.researchgate.net/publication/221334514\\_A\\_Framework\\_for\\_Acoustic\\_Cardiac\\_Signal\\_Analysis](https://www.researchgate.net/publication/221334514_A_Framework_for_Acoustic_Cardiac_Signal_Analysis)>. Citado na página 34.
- CHEN, C. *Signal Processing Handbook*. Taylor & Francis, 1988. (Electrical and Computer Engineering). ISBN 9780824779566. Disponível em: <<https://books.google.com.br/books?id=10Pi0MRbaOYC>>. Citado na página 26.
- DEBBAL, S.; BEREKSI-REGUIG, F. Computerized heart sounds analysis. *Computers in Biology and Medicine*, v. 38, n. 2, p. 263 – 280, 2008. ISSN 0010-4825. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0010482507001771>>. Citado na página 29.
- DELGADO-TREJOS, E. et al. Digital auscultation analysis for heart murmur detection. *Annals of biomedical engineering*, Springer, v. 37, n. 2, p. 337–353, 2009. Citado 3 vezes nas páginas 21, 30 e 35.

- DENG, Y.; BENTLEY, P. J. A robust heart sound segmentation and classification algorithm using wavelet decomposition and spectrogram. In: *Extended Abstract in the First PASCAL Heart Challenge Workshop, held after AISTATS*. [s.n.], 2012. Disponível em: <<http://www.academia.edu/download/30740098/challengepaper3.pdf>>. Citado 3 vezes nas páginas 10, 38 e 46.
- DOKUR, Z.; ÖLMEZ, T. Heart sound classification using wavelet transform and incremental self-organizing map. *Digital Signal Processing*, v. 18, n. 6, p. 951 – 959, 2008. ISSN 1051-2004. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1051200408000961>>. Citado na página 34.
- EL-SEGAIER, M. et al. Computer-based detection and analysis of heart sound and murmur. *Annals of Biomedical Engineering*, Springer, v. 33, n. 7, p. 937–942, 2005. Citado 2 vezes nas páginas 32 e 36.
- ELLIS, D. Chroma feature analysis and synthesis. *Resources of Laboratory for the Recognition and Organization of Speech and Audio-LabROSA*, 2007. Citado na página 25.
- ERGEN, B.; TATAR, Y.; GULCUR, H. O. Time–frequency analysis of phonocardiogram signals using wavelet transform: a comparative study. *Computer methods in biomechanics and biomedical engineering*, Taylor & Francis, v. 15, n. 4, p. 371–381, 2012. Citado na página 31.
- FAWCETT, T. An introduction to roc analysis. *Pattern Recognition Letters*, v. 27, n. 8, p. 861 – 874, 2006. ISSN 0167-8655. ROC Analysis in Pattern Recognition. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S016786550500303X>>. Citado na página 27.
- FRANK, M. A. H. E.; WITTEN, I. H. Data mining: Practical machine learning tools and techniques. The WEKA Workbench. Online Appendix, Morgan Kaufmann, Fourth Edition, 2016. Citado na página 39.
- GEDDES, L. A. Birth of the stethoscope. *IEEE Engineering in Medicine and Biology Magazine*, v. 24, n. 1, p. 84–86, Jan 2005. ISSN 0739-5175. Citado na página 17.
- GHOSH, A. *Mayo Clinic Internal Medicine Review, Eighth Edition*. Taylor & Francis, 2010. (Mayo Clinic Internal Medicine Review). ISBN 9781420084788. Disponível em: <<https://books.google.com.br/books?id=FqVSBol1fcC>>. Citado 3 vezes nas páginas 20, 23 e 35.
- GILL, D.; GAVRIELI, N.; INTRATOR, N. Detection and identification of heart sounds using homomorphic envelopogram and self-organizing probabilistic model. In: *Computers in Cardiology, 2005*. Computers in Cardiology, 2005, 2005. p. 957–960. ISSN 0276-6574. Disponível em: <<https://ieeexplore.ieee.org/document/1588267>>. Citado na página 33.
- GOMES, E. F.; JORGE, A. M.; AZEVEDO, P. J. Classifying heart sounds using multiresolution time series motifs: An exploratory study. In: *Proceedings of the International C\* Conference on Computer Science and Software Engineering*. New York, NY, USA: ACM, 2013. (C3S2E '13), p. 23–30. ISBN 978-1-4503-1976-8. Disponível em: <<http://doi.acm.org/10.1145/2494444.2494458>>. Citado 4 vezes nas páginas 17, 19, 22 e 23.

- GOMES, E. F.; PEREIRA, E. Classifying heart sounds using peak location for segmentation and feature construction. In: . Workshop Classifying Heart Sounds, La Palma, Canary Islands, 2012. Disponível em: <<https://pdfs.semanticscholar.org/62be/8736747856270c757670be80eef659b3c464.pdf>>. Citado 5 vezes nas páginas 10, 13, 14, 38 e 46.
- GOUYON, F.; PACHET, F.; DELERUE, O. On the use of zero-crossing rate for an application of classification of percussive sounds. In: *Proceedings of the COST G-6 conference on Digital Audio Effects (DAFX-00), Verona, Italy*. [s.n.], 2000. Disponível em: <<files/publications/dafx00-gouyon.pdf>>. Citado na página 26.
- HANNA, I. R.; SILVERMAN, M. E. A history of cardiac auscultation and some of its contributors. *The American Journal of Cardiology*, v. 90, n. 3, p. 259 – 267, 2002. ISSN 0002-9149. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0002914902024657>>. Citado 3 vezes nas páginas 19, 20 e 22.
- HARTE, C.; SANDLER, M.; GASSER, M. Detecting harmonic change in musical audio. In: ACM. *Proceedings of the 1st ACM workshop on Audio and music computing multimedia*. 2006. p. 21–26. Disponível em: <<https://dl.acm.org/citation.cfm?id=1178727>>. Citado 2 vezes nas páginas 25 e 26.
- HERMANISKY, H. Perceptual linear predictive (plp) analysis of speech. *the Journal of the Acoustical Society of America, ASA*, v. 87, n. 4, p. 1738–1752, 1990. Citado na página 25.
- JIANG, D.-N. et al. Music type classification by spectral contrast feature. In: IEEE. *Multimedia and Expo, 2002. ICME'02. Proceedings. 2002 IEEE International Conference on*. 2002. v. 1, p. 113–116. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/1035731>>. Citado na página 25.
- KUMAR, D. et al. Heart murmur recognition and segmentation by complexity signatures. In: *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2008. p. 2128–2132. ISSN 1094-687X. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/4649614>>. Citado 2 vezes nas páginas 19 e 35.
- KURNIAWAN, F.; KHALIL, M. S.; MALIK, H. Robust tampered detection method for digital audio using gabor filterbank. *Proc. ICIPCS*, p. 75–82, 2015. Citado na página 25.
- LOGAN, B. et al. Mel frequency cepstral coefficients for music modeling. In: *ISMIR*. ISMIR, 2000. v. 270, p. 1–11. Disponível em: <<http://musicweb.ucsd.edu/~sdubnov/CATbox/Reader/logan00mel.pdf>>. Citado na página 25.
- MARTINEZ-ALAJARIN, R. R.-M. J. *Efficient method for events detection in phonocardiographic signals*. 2005. 5839 - 5839 - 12 p. Disponível em: <<https://doi.org/10.1117/12.608203>>. Citado 2 vezes nas páginas 33 e 35.
- MCFEE, B. et al. *librosa/librosa: 0.6.1*. 2018. Disponível em: <<https://doi.org/10.5281/zenodo.1252297>>. Citado 2 vezes nas páginas 14 e 38.
- MCGILL, H. C.; MCMAHAN, C. A.; GIDDING, S. S. Preventing heart disease in the 21st century: implications of the pathobiological determinants of atherosclerosis in youth (pday) study. *Circulation*, Am Heart Assoc, v. 117, n. 9, p. 1216–1227, 2008. Citado 2 vezes nas páginas 14 e 18.

- MULRONEY SUSAN E MYERS, A. *Netter Bases da Fisiologia 1a edição*. Elsevier Brasil, 2009. ISBN 9788535245868. Disponível em: <<https://books.google.com.br/books?id=t6MIqynoL6AC>>. Citado na página 29.
- NASERI, H.; HOMAEINEZHAD, M. Detection and boundary identification of phonocardiogram sounds using an expert frequency-energy based metric. *Annals of biomedical engineering*, Springer, v. 41, n. 2, p. 279–292, 2013. Citado 4 vezes nas páginas 17, 22, 33 e 35.
- NIEVERGELT, Y. *Wavelets Made Easy*. Birkhäuser Boston, 1999. (Wavelets Made Easy). ISBN 9780817640613. Disponível em: <[https://books.google.com.br/books?id=co5JLG2e\\\_5AC](https://books.google.com.br/books?id=co5JLG2e\_5AC)>. Citado na página 24.
- NIEVERGELT, Y. *Wavelets Made Easy*. Birkhäuser Boston, 2013. ISBN 9781461205739. Disponível em: <<https://books.google.com.br/books?id=AQfpBwAAQBAJ>>. Citado 2 vezes nas páginas 24 e 25.
- NIGAM, V.; PRIEMER, R. Source adaptivity for cardiac sound separation. In: *48th Midwest Symposium on Circuits and Systems, 2005*. IEEE, 2005. p. 1920–1923 Vol. 2. ISSN 1548-3746. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/1594501>>. Citado na página 32.
- ORGANIZATION, W. W. H.; ORGANIZATION, W. H. et al. *The top 10 causes of death*. 2017. (Acessado em 21/05/2018). Citado na página 13.
- OSKIPER, T.; WATROUS, R. Detection of the first heart sound using a time-delay neural network. In: IEEE. *Computers in Cardiology*. 2002. p. 537–540. ISSN 0276-6547. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/1166828>>. Citado na página 33.
- PFAHRINGER, B.; HOLMES, G. *fracpete. Class gridsearch, revision 9733*. 2014. Citado na página 41.
- QUILICI-GONZALEZ, J.; ZAMPIROLI, F. de A. *Sistemas Inteligentes e Mineração de Dados*. Universidade Federal do ABC, 2015. ISBN 9788561175382. Disponível em: <<https://books.google.com.br/books?id=X76VBgAAQBAJ>>. Citado na página 27.
- RABINER, L. R.; JUANG, B.-H. *Fundamentals of speech recognition*. PTR Prentice Hall, 1993. Citado na página 25.
- SLANEY, M. Auditory toolbox. *Interval Research Corporation, Tech. Rep*, v. 10, p. 1998, 1998. Citado na página 25.
- SOARES, H. B. Análise e classificação de imagens de lesões da pele por atributos de cor, forma e textura utilizando máquina de vetor de suporte. Universidade Federal do Rio Grande do Norte, 2008. Citado 2 vezes nas páginas 27 e 38.
- TAVEL, M. E. Cardiac auscultation: a glorious past—and it does have a future! *Circulation*, Am Heart Assoc, v. 113, n. 9, p. 1255–1259, 2006. Citado 4 vezes nas páginas 18, 20, 35 e 36.
- VAPNIK, V. N. *An overview of statistical learning theory*. Wiley, New York, 1999. v. 10. 988-999 p. ISSN 1045-9227. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/788640>>. Citado 3 vezes nas páginas 26, 38 e 41.