



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

**LEANDRO BEZERRA MARINHO**

**SEGMENTAÇÃO DE BATIMENTOS CARDÍACOS EM ELETROCARDIOGRAMA E**  
**IDENTIFICAÇÃO DE ARRITMIA CARDÍACA EM UM ÚNICO LOTE DE**  
**PROCESSAMENTO POR APRENDIZADO PROFUNDO USANDO UMA**  
**ABORDAGEM DE REGRESSÃO EM UMA REDE NEURAL CONVOLUCIONAL**

**FORTALEZA**

**2021**

LEANDRO BEZERRA MARINHO

SEGMENTAÇÃO DE BATIMENTOS CARDÍACOS EM ELETROCARDIOGRAMA E  
IDENTIFICAÇÃO DE ARRITMIA CARDÍACA EM UM ÚNICO LOTE DE  
PROCESSAMENTO POR APRENDIZADO PROFUNDO USANDO UMA ABORDAGEM  
DE REGRESSÃO EM UMA REDE NEURAL CONVOLUCIONAL

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

Orientador: Prof. Dr. Pedro Pedrosa Rebouças Filho

FORTALEZA

2021

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

---

- M29s Marinho, Leandro Bezerra.  
Segmentação de batimentos cardíacos em eletrocardiograma e identificação de arritmia cardíaca em um único lote de processamento por aprendizado profundo usando uma abordagem de regressão em uma rede neural convolucional / Leandro Bezerra Marinho. – 2021.  
111 f. : il. color.
- Tese (doutorado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia de Teleinformática, Fortaleza, 2021.  
Orientação: Prof. Dr. Pedro Pedrosa Rebouças Filho.
1. Eletrocardiograma. 2. Extração de características. 3. Segmentação. 4. Classificação. 5. Aprendizado Profundo. I. Título.

CDD 621.38

---

LEANDRO BEZERRA MARINHO

SEGMENTAÇÃO DE BATIMENTOS CARDÍACOS EM ELETROCARDIOGRAMA E  
IDENTIFICAÇÃO DE ARRITMIA CARDÍACA EM UM ÚNICO LOTE DE  
PROCESSAMENTO POR APRENDIZADO PROFUNDO USANDO UMA ABORDAGEM  
DE REGRESSÃO EM UMA REDE NEURAL CONVOLUCIONAL

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

Aprovada em: 7 de Maio de 2021.

BANCA EXAMINADORA



---

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



---

Prof. Dr. João Paulo Papa  
Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP)



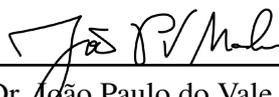
---

Prof. Dr. Frederico Gadelha Guimarães  
Universidade Federal de Minas Gerais (UFMG)



---

Prof. Dr. Paulo Cesar Cortez  
Universidade Federal do Ceará (UFC)



---

Prof. Dr. João Paulo do Vale Madeiro  
Universidade Federal do Ceará (UFC)

À minha esposa, meus pais e irmão pelo incentivo e apoio em todas as minhas escolhas.

## **AGRADECIMENTOS**

A Deus, acima de tudo.

Ao meu orientador, Prof. Dr. Pedro Pedrosa, pelos ensinamentos, amizade, confiança, paciência e dedicação depositados ao longo deste trabalho. Seu conhecimento iluminador e entusiasmo contagiante foram fundamentais para o sucesso desta jornada.

À minha esposa, Renata Lira, pelo carinho, compreensão e incentivo para que eu pudesse concluir esta importante etapa da minha vida.

À minha família que sempre me incentivou a alcançar caminhos cada vez mais distantes.

A todos os meus amigos e companheiros de pesquisa do Laboratório de Processamento de Imagens, Sinais e Computação Aplicada (LAPISCO). Em especial, ao Navar Medeiros, Saulo Oliveira e Solon Alves pelas conversas e discussões que foram de suma importância na conclusão desse trabalho. Além do Jerovane pela ajuda na edição de algumas imagens.

A todos do IFCE que autorizaram o meu limite diferenciado de carga horária. Em especial, aos amigos do campus Acopiara pela compreensão no ajuste da carga horária.

Agradeço a todos os professores por me proporcionar o conhecimento não apenas racional, mas a manifestação do caráter e afetividade da educação no processo de formação profissional, por tanto que se dedicaram a mim, não somente por terem me ensinado, mas por terem me feito aprender.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

“Não haverá borboletas se a vida não passar por  
longas e silenciosas metamorfoses.”

(Rubem Alves)

## RESUMO

Os métodos convencionais para diagnóstico de doenças cardiovasculares em eletrocardiograma (ECG) dependem da obtenção de características particulares no sinal por um observador humano. Devido ao grande número de pacientes e a necessidade de observação contínua, várias técnicas de classificação automática de arritmias foram desenvolvidas nos últimos anos para resolver esse problema. Diante disso, a tese propõe um novo método para classificação de arritmias cardíacas e segmentação de ondas em sinais de eletrocardiograma que usa uma abordagem de regressão com Aprendizado Profundo. Além disso, avalia-se a abordagem clássica de classificação de arritmias que trata de forma independentes as etapas de segmentação, extração de características e classificação. A combinação com maior acurácia, 94,3%, no métodos clássicos foi do extrator Estatísticas de Alta Ordem (*Higher-Order Statistics*, HOS) com o classificador Naive-Bayes usando a divisão com 5 classes. Quanto a problema com 3 classes, os resultados mostram que o HOS obteve desempenho semelhante. Esse extrator em conjunto com o classificador Naive-Bayes conseguiram acurácia de 94%, sensibilidade acima de 99% para os batimentos cardíacos normais e especificidade de 99% nas classes VEB e SVEB. Por conta dos desafios apresentados nesse grupo de estudo, um novo método que usa regressão com Aprendizado Profundo foi proposto. Suas acurácias usando o mesmo protocolo de avaliação para as divisões com 2, 3 e 5 classes foram 99,56%, 99,55% e 99,45%, respectivamente. As outras métricas computadas a partir da matriz de confusão também foram superiores a 99% em todas as variações. O tempo de classificação média do método proposto usando Unidade de Processamento Gráfico (*Graphics Processing Unit*, GPU) foi de 5,59E-3s, corroborando que o mesmo é adequado para sistemas de tempo real. Quanto a detecção do QRS, o método proposto obteve 99,88% de acurácia nas bases de dados MIT-BIH e QT. Em relação a segmentação do QRS, o erro médio da diferença temporal do início do QRS,  $QRS_{on}$ , foi 0,26ms e o  $QRS_{off}$  foi 2,08ms. Na detecção das ondas T e P, as melhores sensibilidades foram 99,85% e 99,52%, enquanto os valores preditivos positivos foram 99,92% e 99,69%. Os erros de segmentação para esses tipos de onda foram  $-0,25 \pm 13,22$  e  $-2,80 \pm 14,79$  para onda T e  $0,26 \pm 9,98$  e  $0,17 \pm 9,68$  para onda P. O estudo apresentado nesta tese mostra que o método proposto mostra-se adequado para integrar sistemas de auxílio ao diagnóstico médico na área de Cardiologia.

**Palavras-chave:** Eletrocardiograma (ECG), Extração de características, Segmentação, Classificação, Aprendizado Profundo

## ABSTRACT

Conventional methods for diagnosing cardiovascular disease on electrocardiogram (ECG) depend on obtaining particular characteristics in the signal by a human observer. Due to many patients and the need for continuous observation, several automatic arrhythmia classification techniques have been developed in recent years to solve this problem. Therefore, the thesis proposes a new method for classifying cardiac arrhythmias and wave segmentation in electrocardiogram signals that uses a Deep Learning regression approach. In addition, we evaluate the classic approach to classifying arrhythmias, which independently deals with the stages of segmentation, feature extraction, and classification. The combination with the highest accuracy, 94.3%, in the classical methods was the extractor *Estatísticas de Alta Ordem (Higher-Order Statistics, HOS)* with the classifier Naive-Bayes using division with five classes. As for the problem with the three classes, the results show that the HOS had a similar performance. This extractor and the Naive-Bayes classifier achieved an accuracy of 94%, sensitivity above 99% for normal heartbeats, and specificity of 99% in the EBV and SVEB classes. Due to the challenges presented in this study group, we propose a new method that uses regression with Deep Learning. Their accuracy using the same evaluation protocol for divisions with 2, 3, and 5 classes was 99.56%, 99.55%, and 99.45%, respectively. The other metrics computed from the confusion matrix were also greater than 99% in all variations. The average classification time of the proposed method using *Unidade de Processamento Gráfico (Graphics Processing Unit, GPU)* was 5.59E-3s, confirming that it is suitable for real-time systems. As for QRS detection, the proposed method obtained 99.88% of accuracy in the MIT-BIH and QT databases. Regarding QRS segmentation, the mean error of the time difference from the beginning of the QRS,  $QRS_{on}$ , was 0.26ms, and the  $QRS_{off}$  was 2.08ms. In detecting T and P waves, the best sensitivities were 99.85% and 99.52%, while the positive predictive values were 99.92% and 99.69%. The segmentation errors for these wave types were  $-0.25 \pm 13.22$  and  $-2.80 \pm 14.79$  for T wave and  $0.26 \pm 9.98$  and  $0.17 \pm 9.68$  for P wave. The study presented in this thesis shows that the proposed method proves to be adequate in integrating systems to aid medical diagnosis in Cardiology.

**Keywords:** Electrocardiogram (ECG), Feature extraction, Segmentation, Classification, Deep learning

## LISTA DE FIGURAS

Figura 1 – Sinal ECG normal e seus eventos característicos. . . . .	16
Figura 2 – Anatomia e sistema de condução do coração. . . . .	31
Figura 3 – Forma esquemática de um período do ECG e amostra de ECG saudável. . .	32
Figura 4 – Disposição tradicional dos eletrodos proposta por Einthoven. . . . .	34
Figura 5 – Disposição do eletrodo torácico proposta por Wilson. . . . .	34
Figura 6 – Disposição dos eletrodos proposta por Goldberger. . . . .	35
Figura 7 – Exemplo de ECG com Fibrilação Atrial. . . . .	36
Figura 8 – ECG de um paciente com arritmia ventricular. . . . .	37
Figura 9 – ECG de um paciente com infarto do miocárdio inferior (a) e ilustração esquemática da elevação do ST durante o ataque cardíaco (b). . . . .	39
Figura 10 – Representação esquemática de um neurônio artificial e gráficos de funções de ativação. . . . .	42
Figura 11 – Função sigmóide (preto) e sua derivada (vermelho). . . . .	42
Figura 12 – Arquitetura de um Perceptron Multicamadas. . . . .	43
Figura 13 – Arquitetura de uma CNN tradicional. . . . .	47
Figura 14 – Ilustração da operação de convolução. A matriz de entrada está destacada em azul, os filtros da camada convolucional em vermelho e o volume de saída resultante da convolução em verde. . . . .	49
Figura 15 – Mapa de ativação gerado por uma camada convolucional com 2 filtros. . . .	50
Figura 16 – Ilustração do efeito da camada de <i>pooling</i> com operador máximo. . . . .	51
Figura 17 – Fluxograma das etapas adotadas na abordagem clássica para classificar arritmias cardíacas em sinais de ECG. . . . .	53
Figura 18 – Etapas adotadas no método para obtenção do diagnóstico médico usando sinal do eletrocardiograma. . . . .	58
Figura 19 – Trechos aleatórios do registro 100 nas base de dados MIT-BIH e QT. . . . .	60
Figura 20 – Exemplo da criação da matriz $\mathbf{M}$ após a transformação $\tau(\cdot)$ . . . . .	61
Figura 21 – (a) exemplo da $K$ regiões criadas após a transformação $\tau(\cdot)$ e (b) métrica empregada para calcular a precisão de segmentação dos batimentos cardíacos, $j(O_{real}, O_{predita})$ . . . . .	63
Figura 22 – Diagrama da arquitetura unificada para segmentação e classificação do batimentos cardíacos. . . . .	64

Figura 23 – Arquitetura da rede neural convolucional adotada para segmentar e classificar simultaneamente os batimentos cardíacos do sinal ECG. . . . .	65
Figura 24 – Duas formas de visualizar a classificação dos batimentos cardíacos. (a) exemplo do mapa de probabilidades para os tipos de batimento cardíaco e (b) $J$ janelas criadas por regiões. . . . .	67
Figura 25 – Predições antes e depois do algoritmo de seleção de janela sobre cada tipo de batimento. . . . .	69
Figura 26 – Exemplo da matriz de confusão para um problema multiclasse. . . . .	75
Figura 27 – Comparação entre os métodos da abordagem clássica proposta nesta tese e dos trabalhos de Luz <i>et al.</i> (2013). A melhor abordagem é destacada pela barra hachurada. . . . .	85
Figura 28 – Comparação entre o histograma do registro de ECG para o paciente 101 e uma distribuição Gaussiana. . . . .	86
Figura 29 – Resultado do método proposto por classe para classificação binária de arritmias cardíacas no ECG usando diferentes paradigmas de avaliação. . . . .	88
Figura 30 – Resultado do método proposto por classe na classificação de arritmias cardíacas usando a divisão em 3 classes. . . . .	89
Figura 31 – Resultado do método proposto por classe na classificação de arritmias cardíacas usando a divisão em 5 classes. . . . .	89
Figura 32 – Resultados da saída do método proposto para intervalos de dois registros da base de dados QT. . . . .	94

## LISTA DE TABELAS

Tabela 1 – Exemplos de funções de ativação . . . . .	41
Tabela 2 – Agrupamento das anotações da MIT-BIH segundo a norma ANSI/AAMI EC57.	72
Tabela 3 – Conjuntos de dados propostos por Chazal <i>et al.</i> (2004) para treinar (DS1) e avaliar os modelos (DS2). . . . .	72
Tabela 4 – Quantidade de atributos considerando a divisão com 5 classes. . . . .	73
Tabela 5 – Quantidade de atributos considerando a divisão com 3 classes. . . . .	73
Tabela 6 – Resumo das métricas adotadas na classificação das arritmias cardíacas. . . .	75
Tabela 7 – Conjunto de parâmetros usado pelo método proposto nesta tese. . . . .	76
Tabela 8 – Métodos da abordagem clássica que usam o paradigma inter-paciente. . . .	78
Tabela 9 – Métodos da abordagem clássica que usam o paradigma intra-paciente. . . .	79
Tabela 10 – Resultados das combinações de extratores e classificadores na base de dados MIT-BIH considerando a divisão de 5 classes com a abordagem clássica. . . .	81
Tabela 11 – Resultados das combinações de extratores e classificadores na base de dados MIT-BIH considerando a divisão de 3 classes com a abordagem clássica. . . .	83
Tabela 12 – Resultados do tempo de treinamento, teste e de extração considerando a divisão de 3 classes. . . . .	84
Tabela 13 – Resultados do método proposto para classificação das arritmias cardíacas considerando diferentes divisões de classes e paradigmas de avaliação. . . .	86
Tabela 14 – Comparação entre os melhores resultados obtidos pelo método proposto e os trabalhos relacionados usando aprendizado profundo para classificar arritmia cardíaca. . . . .	87
Tabela 15 – Custo computacional para treinar o método e predizer um batimento cardíaco considerando diferentes divisões de classe e paradigmas de avaliação. . . . .	90
Tabela 16 – Resultados da detecção do complexo QRS sobre os bancos de dados MIT-BIH e QT. <i>An</i> indica a quantidade de anotações médicas usadas na avaliação. . . .	91
Tabela 17 – Média e desvio padrão da diferença temporal, em <i>ms</i> , para a segmentação do QRS sobre o banco de dados QT. . . . .	91
Tabela 18 – Resultados do método proposto para detecção e segmentação das ondas T e P no banco de dados QT. . . . .	92
Tabela 19 – Resultado da diferença entre os limites da onda real e a predita na matriz <b>M</b> sobre a base de dados QT. . . . .	93

## LISTA DE ALGORITMOS

Algoritmo 1 – Algoritmo de seleção das janelas preditas pelo modelo. . . . .	69
--	----

## LISTA DE ABREVIATURAS E SIGLAS

AAMI	<i>Advancement of Medical Instrumentation</i>
AF	Fibrilação Atrial
ANN	<i>Artificial Neural Network</i> (Rede Neural Artificial)
CAD	<i>Computer-aided diagnosis</i> (Diagnóstico Auxiliado por Computador)
CNN	<i>Convolutional Neural Network</i> (Rede Neural Convolucional)
DNN	<i>Deep Neural Network</i> (Rede Neural Profunda)
DWT	<i>Discrete Wavelet Transform</i> (Transformada Wavelet Discreta)
ECG	Eletrocardiograma
FFT	<i>Fast Fourier Transform</i> (Transformada Rápida de Fourier)
GPU	<i>Graphics Processing Unit</i> (Unidade de Processamento Gráfico)
HOS	<i>Higher-Order Statistics</i> (Estatísticas de Alta Ordem)
ICA	<i>Independent Component Analysis</i> (Análise de Componentes Independentes)
LDA	<i>Linear Discriminant Analysis</i> (Análise de Discriminantes Lineares)
LSTM	<i>Long Short-Term Memory</i> (Memória Longa de Curto Prazo)
MLP	<i>Multilayer Perceptron</i> (Perceptron Multicamada)
OPF	<i>Optimum-Path Forest</i> (Floresta de Caminhos Ótimos)
PCA	<i>Principal Component Analysis</i> (Análise de Componentes Principais)
RBF	<i>Radial Basis Function</i> (Função de Base Radial)
RNN	<i>Recurrent Neural Network</i> (Rede Neural Recorrente)
SCM	<i>Structural Co-occurrence Matrix</i> (Matriz de Co-ocorrência Estrutural)
STFT	<i>Short-Time Fourier Transform</i> (Transformada de Fourier de Curta Duração)
SVM	<i>Support Vector Machine</i> (Máquina de Vetores de Suporte)
TV	Taquicardia Ventricular
VCG	Vetorcardiograma

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>16</b>
<b>1.1</b>	<b>Motivação</b>	<b>18</b>
<b>1.2</b>	<b>Estado da arte</b>	<b>19</b>
<b>1.3</b>	<b>Objetivos</b>	<b>23</b>
<i>1.3.1</i>	<i>Objetivos específicos</i>	<i>23</i>
<b>1.4</b>	<b>Contribuições</b>	<b>24</b>
<b>1.5</b>	<b>Produção Científica</b>	<b>24</b>
<i>1.5.1</i>	<i>Artigos completos publicados em periódicos</i>	<i>25</i>
<i>1.5.2</i>	<i>Trabalhos completos publicados em anais de congressos</i>	<i>26</i>
<i>1.5.3</i>	<i>Capítulos de livros publicados</i>	<i>27</i>
<i>1.5.4</i>	<i>Artigos em processo de avaliação</i>	<i>27</i>
<i>1.5.5</i>	<i>Registros de softwares</i>	<i>28</i>
<b>1.6</b>	<b>Organização da tese</b>	<b>28</b>
<b>2</b>	<b>FUNDAMENTOS TEÓRICOS</b>	<b>30</b>
<b>2.1</b>	<b>Aspectos do coração e eletrocardiograma</b>	<b>30</b>
<i>2.1.1</i>	<i>Funções e estrutura do coração</i>	<i>30</i>
<i>2.1.2</i>	<i>Sinal eletrocardiograma</i>	<i>32</i>
<i>2.1.3</i>	<i>Disposição dos eletrodos para o eletrocardiograma</i>	<i>33</i>
<i>2.1.4</i>	<i>Doenças cardiovasculares e eletrocardiograma</i>	<i>35</i>
<b>2.2</b>	<b>Introdução ao Aprendizado Profundo</b>	<b>40</b>
<b>2.3</b>	<b>Redes Neurais Artificiais</b>	<b>40</b>
<i>2.3.1</i>	<i>Neurônio Artificial</i>	<i>41</i>
<i>2.3.2</i>	<i>Perceptron Multicamadas</i>	<i>42</i>
<b>2.4</b>	<b>Redes Neurais Convolucionais</b>	<b>46</b>
<i>2.4.1</i>	<i>Camada convolucional</i>	<i>47</i>
<i>2.4.2</i>	<i>Camada de agrupamento</i>	<i>50</i>
<i>2.4.3</i>	<i>Camada totalmente conectada</i>	<i>51</i>
<b>2.5</b>	<b>Considerações finais</b>	<b>51</b>
<b>3</b>	<b>METODOLOGIA</b>	<b>52</b>

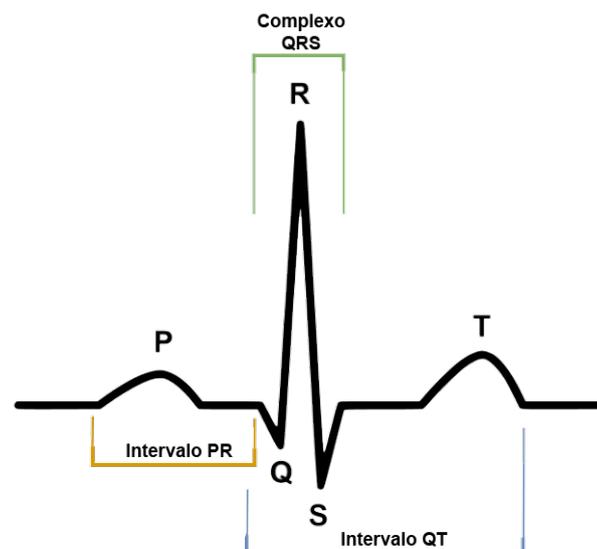
<b>3.1</b>	<b>Abordagem clássica para classificação de arritmias cardíacas usando sinais de ECG . . . . .</b>	<b>52</b>
<b>3.1.1</b>	<i>Métodos de extração de características . . . . .</i>	<b>54</b>
<b>3.1.2</b>	<i>Parâmetros para segmentar os batimentos na abordagem clássica . . . . .</i>	<b>55</b>
<b>3.1.3</b>	<i>Métodos de classificação . . . . .</i>	<b>56</b>
<b>3.2</b>	<b>Formulação matemática do método proposto . . . . .</b>	<b>57</b>
<b>3.2.1</b>	<i>Pré-processamento do sinal de entrada . . . . .</i>	<b>59</b>
<b>3.2.2</b>	<i>Processamento das janelas usando regressão com Aprendizado Profundo .</i>	<b>61</b>
<b>3.2.2.1</b>	<i>Projeto e treinamento da Rede Neural Convolutacional . . . . .</i>	<b>64</b>
<b>3.2.2.2</b>	<i>Junção das bases de segmentação e classificação durante o treinamento . .</i>	<b>68</b>
<b>3.2.3</b>	<i>Pós-processamento . . . . .</i>	<b>68</b>
<b>3.3</b>	<b>Especificações para caracterização das arritmias cardíacas no eletrocardiograma . . . . .</b>	<b>70</b>
<b>3.3.1</b>	<i>Descrição da base MIT-BIH . . . . .</i>	<b>70</b>
<b>3.3.2</b>	<i>Agrupamento dos batimentos cardíacos e atributos dos extratores de características . . . . .</i>	<b>71</b>
<b>3.4</b>	<b>Base de dados para segmentação e detecção das ondas no ECG . . . . .</b>	<b>72</b>
<b>3.5</b>	<b>Métricas de avaliação . . . . .</b>	<b>74</b>
<b>3.6</b>	<b>Considerações finais . . . . .</b>	<b>76</b>
<b>4</b>	<b>RESULTADOS . . . . .</b>	<b>77</b>
<b>4.1</b>	<b>Resultados da classificação de arritmia cardíaca com a abordagem clássica</b>	<b>77</b>
<b>4.1.1</b>	<i>Resultado dos métodos clássicos utilizando o protocolo inter-paciente . . .</i>	<b>80</b>
<b>4.2</b>	<b>Resultados da classificação de arritmia cardíaca usando Aprendizado Profundo . . . . .</b>	<b>85</b>
<b>4.3</b>	<b>Resultados da detecção e segmentação dos batimentos no ECG . . . . .</b>	<b>89</b>
<b>4.4</b>	<b>Considerações finais . . . . .</b>	<b>92</b>
<b>5</b>	<b>CONCLUSÕES E TRABALHOS FUTUROS . . . . .</b>	<b>97</b>
	<b>REFERÊNCIAS . . . . .</b>	<b>101</b>

## 1 INTRODUÇÃO

As doenças cardiovasculares são a principal causa de óbito no mundo. Estima-se que 17,9 milhões de pessoas morreram dessas doenças em 2016 (WHO, 2018). Por conta do baixo custo e devido as características não invasivas, o Eletrocardiograma (ECG) é o principal exame para o diagnóstico de anomalias cardíacas. A análise visual do ECG, no entanto, é imprecisa devido à quantidade de informações. Assim, o desenvolvimento de sistemas que auxiliem no diagnóstico médico é imprescindível (ZHOU; TAN, 2020).

Os eventos decorrentes da despolarização e polarização do tecido cardíaco constituem o sinal do ECG. As ondas mais relevantes são as P, QRS e T e representam a despolarização atrial, despolarização ventral e repolarização (MARINHO *et al.*, 2019). A Figura 1 ilustra um batimento cardíaco normal em sinal de ECG com suas ondas e alguns intervalos.

Figura 1 – Sinal ECG normal e seus eventos característicos.



Fonte: Adaptado de Marinho *et al.* (2019).

As arritmias cardíacas são causadas por alterações na formação e condução dos impulsos elétricos pelo tecido miocárdico, que podem modificar a origem e a difusão fisiológica do estímulo elétrico do coração. Também chamadas de disritmias cardíacas, elas indicam um distúrbio do ritmo cardíaco, que pode variar em gravidade desde totalmente benigno até arritmias com risco de vida imediato. Esses batimentos atípicos estão fortemente ligados a patologias cardíacas, como o infarto do miocárdio (LI; BOULANGER, 2020). Portanto, a detecção e caracterização das arritmias para o diagnóstico cardíaco é de fundamental importância.

Os sistemas convencionais de classificação de anomalias cardíacas a partir de sinais

de ECG são divididos nas seguintes etapas: aquisição, pré-processamento, segmentação, extração de características e classificação (NASCIMENTO *et al.*, 2020). Muitas das informações que podem ser clinicamente relevantes não são evidentes durante a análise visual do sinal de ECG, exigindo assim métodos específicos de extração de características. As abordagens mais comuns utilizam a Transformada Wavelet, Análise de Componentes Principais e espectro de potência (LUZ *et al.*, 2016).

As pesquisas da literatura que seguem as etapas supracitadas são divididas em dois tipos principais de paradigmas de avaliação. No primeiro, os trabalhos seguem o paradigma intra-paciente, onde os batimentos cardíacos dos mesmos pacientes provavelmente aparecem tanto no treinamento quanto no conjunto de dados de teste. Enquanto no paradigma inter-paciente, os autores constroem e avaliam a classificação utilizando os batimentos cardíacos de diferentes pacientes, seguindo o protocolo proposto por Chazal *et al.* (2004).

O uso de métodos de Aprendizado Profundo (*Deep Learning*) para classificação de arritmias cardíacas em sinais de ECG é cada vez mais recorrente. Além do excelente desempenho para prever padrões, essas técnicas não demandam explicitamente a etapa de extração de características. As abordagens de aprendizagem profunda para essa tarefa costumam utilizar Rede Neural Profunda (*Deep Neural Network*, DNN), representação da imagem espectral, entropia espectral, Rede Neural Convolucional (*Convolutional Neural Network*, CNN) e Memória Longa de Curto Prazo (*Long Short-Term Memory*, LSTM) (MURAT *et al.*, 2020).

Esta teste tem como principal contribuição a criação de um modelo baseado em Aprendizado Profundo que visa segmentar os batimentos e ao mesmo tempo classificar as arritmias cardíacas. Ou seja, busca-se resposta para a seguinte pergunta. **É possível obter bons resultados usando Aprendizado Profundo para análise de sinais de ECG tratando o problema como regressão?** Além disso, o método proposto foi comparado com estratégias clássicas que segmenta os batimentos e classifica as arritmias de forma independente. Por conta dos desafios da estratégia citada anteriormente, um novo método que usa regressão com Aprendizado Profundo é proposto. Além dos excelentes resultados na classificação, o método é capaz de simultaneamente detectar e segmentar os batimentos cardíacos em sinais de ECG. Ademais, o método também extrai os atributos relevantes para classificação das anomalias cardíacas sem a intervenção de um especialista. Neste sentido, a técnica proposta é comparada com outras de segmentação e classificação para acelerar e melhorar a análise de patologias cardíacas em sinais de eletrocardiograma.

## 1.1 Motivação

A Organização Mundial da Saúde (OMS) estima que em 2016 as doenças cardiovasculares foram responsáveis por 31% de todas as mortes no mundo. Dessas mortes, 4 em cada 5 mortes são devido a ataques cardíacos e derrames, e um terço delas ocorrem prematuramente em pessoas com menos de 70 anos de idade (WHO, 2018).

Indivíduos com risco de doenças cardiovasculares podem apresentar aumento da pressão arterial, glicose e lipídios, bem como sobrepeso e obesidade. Tudo isso pode ser facilmente medido em unidades de atenção primária. Identificar aqueles com maior risco dessas doenças e garantir que recebam o tratamento adequado pode prevenir mortes prematuras (WHO, 2018).

O processo de identificação e classificação de anomalias cardíacas é incômodo ao ser humano, pois na maioria dos casos é necessário analisar cada batimento dos registros do ECG, adquiridos de um monitor *holter* durante horas ou mesmo dias. Além disso, existe a possibilidade de erro humano durante a análise dos registros do ECG devido à fadiga. Uma alternativa é usar técnicas computacionais para classificação automática (BERAZA; ROMERO, 2017).

Os avanços no Aprendizado de Máquina proporcionaram o desenvolvimento eficiente de sistemas de Diagnóstico Auxiliado por Computador (*Computer-aided diagnosis*, CAD) e sua aplicação em muitas áreas. O surgimento desses sistemas juntamente com a precisão na detecção de doenças permitiu a diminuição da carga de trabalho dos cardiologistas. Para minimizar os erros visuais e compensar a interpretação manual, os pesquisadores começaram a desenvolver sistemas CAD para auxiliar no diagnóstico por meio do ECG (LUZ *et al.*, 2016).

Um sistema CAD eficaz requer um classificador de padrões poderoso, bem como um extrator de características relevante para obter informações significativas dos dados brutos. Os atributos utilizados por esses métodos são geralmente projetados e selecionados por tentativa e erro ou experiência. Portanto, esses sistemas exigem conhecimentos mais específicos em vários domínios tornando o processo todo mais complexo. As técnicas de Aprendizado Profundo foram desenvolvidas para superar essas dificuldades e fornecer taxas de detecção aprimoradas sem o uso de atributos codificados de maneira fixa (MURAT *et al.*, 2020).

Os métodos tradicionais de rede neural artificial e classificadores que se baseiam em *kernel* estão entre os mais usados para classificação de anomalias cardíacas. Essas técnicas geralmente usam entrada obtidas por extratores de características no lugar dos sinais de entrada. Uma vez que o alto desempenho desejado dos dados de entrada não pôde ser alcançado, as

pesquisas anteriores focaram nos métodos de extração de características ao invés da estrutura das redes neurais artificiais. O tempo, frequência, propriedades estatísticas e não lineares são obtidos por abordagens como a transformada Wavelet, transformada de Fourier e Estatísticas de Alta Ordem (*Higher-Order Statistics*, HOS) (MARINHO *et al.*, 2019). Devido à alta dimensionalidade dos vetores de características resultantes de métodos de transformação, o tamanho desses vetores é reduzido com abordagens ou técnicas estatísticas, como Análise das Componentes Principais (LUZ *et al.*, 2016).

As técnicas de Aprendizado Profundo, que são cada vez mais populares na literatura, fornecem um meio eficaz para a extração de informação sem a necessidade de engenharia de atributos. Ademais, as estruturas desses métodos têm o potencial de aprender os atributos obtidos por especialistas de domínios, bem como características desconhecidas (LECUN *et al.*, 2015).

Neste contexto, um novo método para classificação de arritmias cardíacas e segmentação de ondas em sinais de eletrocardiograma que usa uma abordagem de regressão com Aprendizado Profundo é proposto e avaliado. Além disso, investiga-se a abordagem clássica de classificação de arritmias que trata de forma independentes as etapas de segmentação, extração de características e classificação. Nessa abordagem, diferentes técnicas de extração e classificação foram analisadas. Diante dos resultados obtidos, motivou-se a criação do método proposto com uso de Aprendizado Profundo. Ao contrário das técnicas tradicionais de Processamento Digital de Sinais e Aprendizado de Máquina, o método proposto usa um único modelo para realizar a segmentação, detecção e classificação simultaneamente. Outro diferencial do modelo é sua capacidade de extrair características sem a intervenção de um especialista e de maneira adaptativa.

## 1.2 Estado da arte

O ECG é o sinal biomédico mais adotado em diagnósticos de doenças cardiovasculares. Suas derivações, que capturam o potencial elétrico do coração de diferentes ângulos e posições, podem ser usadas para indicar o estado da doença por meio de anormalidades nas formas de onda ou ritmos (NAJARIAN; SPLINTER, 2012). Portanto, é de fundamental importância a detecção e caracterização desses batimentos para o diagnóstico cardíaco.

A detecção e classificação automática de arritmias em eletrocardiogramas é amplamente estudada e empregada. A literatura apresenta vários métodos computacionais desde a etapa de segmentação até a classificação. A maioria deles aborda a detecção com um número

limitado de segmentos do ECG. Usar o mesmo método para detectar todos os eventos pode ser mais desafiador devido às diferentes características entre as ondas no ECG em termos de forma, frequência, amplitude e duração (LUZ *et al.*, 2016).

A maioria dos algoritmos de segmentação adotam duas etapas. Na primeira, o sinal de ECG é transformado ou filtrado para destacar a seção relevante do ECG. Isso geralmente é feito aplicando vários métodos de filtragem ou transformação, como transformada Wavelet (MARTINEZ *et al.*, 2004; GHAFARI *et al.*, 2009; MOURAD; FETHI, 2016; LI *et al.*, 1995), transformada de Hilbert (MADEIRO *et al.*, 2012), demodulação complexa de frequência (BASHAR *et al.*, 2019), filtragem linear e não linear (HAMILTON; TOMPKINS, 1986), filtro diferencial passa-baixa (LAGUNA *et al.*, 1994) e transformada Fasorial (MARTINEZ *et al.*, 2010).

Na segunda etapa, o algoritmo identifica e localiza os pontos de interesse definindo um conjunto de regras. Isso é comumente alcançado pelo uso de uma abordagem heurística que se baseiam em um limiar adaptativo (QIN *et al.*, 2017). Alternativamente, alguns algoritmos aplicam modelos estocásticos ou probabilísticos, como o modelo oculto de Markov (ANDREÃO; BOUDY, 2006), Redes Neurais Artificiais (KANG *et al.*, 2015), algoritmos evolutivos (DUMONT *et al.*, 2008) e técnicas Bayesianas (LIN *et al.*, 2010).

A etapa de extração de características é uma das mais importantes nas abordagens tradicionais para a classificação eficiente dos batimentos cardíacos usando o sinal de ECG. Os atributos podem ser extraídas de várias formas, diretamente da morfologia do sinal do ECG, no domínio do tempo, da frequência ou do ritmo cardíaco (LUZ *et al.*, 2016).

Os atributos mais comuns encontrados na literatura são calculados a partir do intervalo dos batimentos cardíacos, também conhecido como intervalo RR. Com exceção dos pacientes que utilizam marca-passo, as variações percebidas na largura do intervalo RR se correlacionam com as da morfologia da onda, frequentemente provocadas por arritmias. Outras pesquisas utilizaram a duração do complexo QRS como atributo (CLIFFORD *et al.*, 2006).

Os atributos extraídos no domínio de tempo e da frequência junto com os do intervalo RR aparecem como parte dos métodos que produziram as maiores precisões na literatura usando as abordagens tradicionais. A maneira mais simples de extrair características no domínio do tempo é utilizar os pontos da curva segmentada do ECG (WEN *et al.*, 2009). Essa abordagem, no entanto, gera vetores com de alta dimensionalidade e sofrem problemas relacionados à escala ou deslocamento do sinal em relação ao pico R (ÖZBAY; TEZEL, 2010).

A transformada Wavelet está presente em várias pesquisas da literatura e os autores afirmam ser um dos melhores métodos para extrair características do sinal de ECG. A transformada Wavelet permite a extração de informações tanto do domínio da frequência quanto do domínio do tempo, diferente do que normalmente é obtido pela transformada de Fourier (KUTLU; KUNTALP, 2012). A escolha da função Wavelet mãe, a ordem dos filtros e o nível de decomposição escolhidos para extração são cruciais para o desempenho final do modelo de classificação. Esses parâmetros devem ser analisados cuidadosamente para não perder detalhes importantes do sinal de ECG (GÜLER; ÜBEYLI, 2005).

Dentre os diversos extratores existentes na literatura, alguns podem ser destacados pelo bom desempenho. A Transformada de Fourier é uma poderosa ferramenta capaz de destacar características relevantes de um sinal no domínio da frequência (ERSOY, 1994). Pławiak (2018) propôs uma metodologia para classificação de doenças cardíacas em sinais de ECG usando transformada de Fourier. Goertzel (1958) desenvolveu um algoritmo para otimizar o cálculo dos coeficientes discretos da transformada de Fourier. Outro extrator importante usado em sinais unidimensionais é a Estatísticas de Alta Ordem (*Higher-Order Statistics*, HOS). Este método possui bom desempenho em problemas onde a não Gaussianidade, ruído colorido ou não linearidades são essenciais e devem ser levados em consideração (MENDEL, 1991). Assim, o HOS pode ser um método adequado para extrair características em sinal de ECG. A Matriz de Co-ocorrência Estrutural (*Structural Co-occurrence Matrix*, SCM) é um extrator de características invariante à rotação que se baseia em uma abordagem estrutural usando matriz de co-ocorrência. Essa matriz armazena informações sobre as dissimilaridades das estruturas de dois sinais de entrada (RAMALHO *et al.*, 2016). Souza *et al.* (2018) propuseram uma abordagem para diagnosticar a doença de Parkinson usando SCM para análise de similaridade.

As pesquisas que não usam o paradigma inter-paciente dificultam as comparações dos resultados na literatura, pois a grande maioria não segue o mesmo protocolo de avaliação. Além disso, estes trabalhos podem não ser clinicamente relevantes por apresentar resultados que podem divergir do contexto real. Diante disso, Luz *et al.* (2013) reavaliaram alguns extratores de características com taxa de acerto próxima de 100%, mas que não seguiam nenhum protocolo de seleção de batimentos cardíacos. Estes métodos baseiam-se na Transformada Wavelet Discreta (*Discrete Wavelet Transform*, DWT), Análise de Componentes Independentes (*Independent Component Analysis*, ICA), Análise de Componentes Principais (*Principal Component Analysis*, PCA) e informações sobre o intervalo RR.

Por fim, os modelos para classificar as anomalias cardíacas são construídos, a partir dos conjuntos de atributos, usando algoritmos de Aprendizado de Máquina. Os métodos mais populares empregados para esta tarefa encontrados na literatura são a Máquina de Vetores de Suporte (*Support Vector Machine*, SVM), Rede Neural Artificial (*Artificial Neural Network*, ANN) e Análise de Discriminantes Lineares (*Linear Discriminant Analysis*, LDA) (LUZ *et al.*, 2016).

A principal desvantagem da abordagem tradicional para classificação de anomalias cardíacas em sinais ECG é a alta complexidade para modelar os métodos responsáveis por cada uma das etapas descritas anteriormente. As técnicas de Aprendizado Profundo, por outro lado, superam essas limitações ao mesmo tempo que aprimoram os resultados dos diagnósticos (MURAT *et al.*, 2020). Os métodos de aprendizagem profunda para classificação de anomalias cardíacas em sinais de ECG costumam adotar Rede Neural Profunda (*Deep Neural Network*, DNN) (XU *et al.*, 2019; SANNINO; PIETRO, 2018; LUO *et al.*, 2017), Rede Neural Convolutiva (*Convolutional Neural Network*, CNN) (LI *et al.*, 2018; YILDIRIM *et al.*, 2018; ZHAI; TIN, 2018; KIRANYAZ *et al.*, 2016), Transformada de Fourier de Curta Duração (*Short-Time Fourier Transform*, STFT) com CNN (HUANG *et al.*, 2019), autoencoder convolutiva (OH *et al.*, 2019) e CNN com Memória Longa de Curto Prazo (*Long Short-Term Memory*, LSTM) (OH *et al.*, 2018).

Neste sentido, esta tese propõe a criação de um novo método capaz de simultaneamente realizar a classificação das doenças cardiovasculares, detectar e segmentar os batimentos cardíacos em sinais de ECG. Através de uma abordagem de regressão com Aprendizado Profundo, o método resolve de forma adaptativa e sem intervenção humana uma das etapas mais críticas na classificação das anomalias cardíacas em sinais de ECG que é a extração de características. O método proposto é avaliado e comparado com outras técnicas e diferentes bases de dados quanto à qualidade de segmentação, detecção e classificação anomalias cardíacas em sinais de ECG. Além disso, é feito um estudo de diferentes extratores e classificadores clássicos para classificação de arritmias cardíacas em sinais de eletrocardiograma que tratam essas etapas de maneira independente. Os desafios presentes neste último cenário de investigação motivou a criação do método proposto.

### 1.3 Objetivos

O principal objetivo desta tese é desenvolver um método que usa uma abordagem de regressão com Aprendizado Profundo capaz de simultaneamente classificar arritmias cardíacas, detectar e segmentar os batimentos em sinais de eletrocardiograma.

#### 1.3.1 *Objetivos específicos*

Os objetivos específicos que devem ser alcançados neste trabalho são:

- fornecer uma análise ampla de métodos de extração de características e classificação que usam abordagem clássica para classificação da arritmia cardíacas em sinais ECG;
- avaliar como a combinação da SCM e Fourier se comportará em sinais de ECG, uma vez que só foi aplicada em imagens;
- propor um método que use regressão com Aprendizado Profundo para simultaneamente classificar arritmias cardíacas, segmentar e detectar os batimentos do coração em sinais de eletrocardiograma;
- utilizar uma abordagem de regressão com Aprendizado Profundo para extrair características de forma adaptativa e sem intervenção humana para classificação de arritmias cardíacas;
- avaliar a classificação de arritmias cardíacas em sinais de ECG seguindo as recomendações ANSI/AAMI EC57:1998/(R) 2008;
- agrupar e analisar as anotações médicas sobre batimento cardíaco em cinco classes para classificação de arritmias;
- utilizar e avaliar o agrupamento das anotações médicas em N, SVEB e VEB para classificação de arritmias;
- propor e avaliar o método em uma abordagem binária que agrupa as classes em normal ou com arritmia cardíaca;
- dividir os registros em dois conjuntos de dados para treinamento e validação do modelo utilizando os paradigmas intra-paciente, inter-paciente e com todo o conjunto de dados;
- analisar a eficiência do método para detecção do complexo QRS nas bases de dados MITDB e QTDB;
- avaliar o método no problema de segmentação e detecção das ondas T, P e QRS e comparar com outros trabalhos do estado da arte;

- propor e avaliar métricas para a matriz de transformação **M**;
- produzir trabalhos científicos e/ou tecnológicos.

#### 1.4 Contribuições

As principais contribuições desta tese são:

- Análise dos impactos, do ponto de vista clínico, de combinações de extratores e classificadores que seguem uma abordagem clássica na classificação de arritmias cardíacas.
- Elaboração de método automático para classificar anomalias cardíacas e, ao mesmo tempo, segmentar e detectar os batimentos em sinais de eletrocardiograma.
- Desenvolvimento de uma abordagem de regressão com Aprendizado Profundo para auxílio de diagnóstico médico capaz de classificar doenças cardiovasculares e segmentar os batimentos no ECG simultaneamente.
- Proposição de uma transformação no sinal de entrada para permitir a combinação de diferentes canais de derivação em uma Rede Neural Convolutacional.
- Avaliação do problema utilizando o agrupamento das anotações médicas em 5, 3 e 2 classes. A primeira divisão segue a norma ANSI/AAMI EC57:1998/(R) 2008. A quantidade não representativa de amostras da classe Q e os valores inexpressivos alcançados pela F levaram a divisão do problema em três classes principais. Juntamente com essas divisões, e ao contrário da maioria dos trabalhos na literatura, também é realizado o estudo binário do problema.
- Divisão e análise dos registros seguindo os paradigmas intra-paciente, inter-paciente e com todo o conjunto de dados. Diversas pesquisas seguem o paradigma intra-paciente, mas são menos relevante do ponto de vista clínico.
- Avaliação da eficiência do método para segmentar e detectar as ondas T, P e QRS nas bases MITDB e QTDB. A segmentação é realizada concomitantemente a classificação das anomalias cardíacas por um único modelo.

#### 1.5 Produção Científica

Esta seção apresenta as produções científicas realizadas no período de desenvolvimento desta pesquisa. A Seção 1.5.1 apresenta os artigos completos publicados em periódicos, a Seção 1.5.2 mostra os trabalhos completos publicados em anais de congresso, a Seção 1.5.3 traz

os capítulos de livros publicados, enquanto os artigos em processo de avaliação são descritos na Seção 1.5.4. Por fim, a Seção 1.5.5 mostra os registros de software.

Os artigos ligados diretamente a tese são os dois no início da Seção 1.5.1 e o primeiro da Seção 1.5.4. Os demais trabalhos utilizam as ferramentas que ajudaram na construção da mesma.

### ***1.5.1 Artigos completos publicados em periódicos***

- **MARINHO, L. B.**; NASCIMENTO, N. DE M.M.; SOUZA, J. W. M. ; GURGEL, M. V.; REBOUÇAS F., P. P.; DE ALBUQUERQUE, V. H. C. A novel electrocardiogram feature extraction approach for cardiac arrhythmia classification. *Future Generation Computer Systems*, v. 97, p. 564-577, 2019.
- NASCIMENTO, N. M. M.; **MARINHO, L. B.**; PEIXOTO, S. A.; DO VALE ., J. P.; DE ALBUQUERQUE, V. H. C.; REBOUÇAS F., P. P. Heart Arrhythmia Classification Based on Statistical Moments and Structural Co-occurrence. *Circuits Systems and Signal Processing*, v. 38, p. 1-20, 2019.
- **MARINHO, L. B.**; REBOUÇAS F., P. P.; DE ALBUQUERQUE, V. H. C. Ultrasonic sensor signals and self organized mapping with nearest neighbors for the microstructural characterization of thermally-aged Inconel 625 alloy. *Computers in Industry*, v. 107, p. 1-10, 2019.
- ALMEIDA, J. S.; **MARINHO, L. B.**; SOUSA, J. W. M.; ASSIS, E. A.; REBOUCAS F., P. P. Localization System for Autonomous Mobile Robots Using Machine Learning Methods and Omnidirectional Sonar. *IEEE Latin America Transactions*, v. 16, p. 368-374, 2018.
- HOLANDA, G. B.; SOUZA, J. W. M.; LIMA, DANIEL A.; **MARINHO, L. B.**; GIRÃO, A. M.; BEZERRA F., J. B.; REBOUÇAS F., P. P. Development of OCR system on android platforms to aid reading with a refreshable braille display in real time. *Measurement*, v. 120, p. 150-168, 2018.
- **MARINHO, L. B.**; REBOUÇAS F., P. P.; ALMEIDA, J. S.; SOUZA, J. W. M.; SOUZA J., A. H.; DE ALBUQUERQUE, V. H. C.. A novel mobile robot localization approach based on classification with rejection option using computer vision. *Computers & Electrical Engineering*, v. 68, p. 26-43, 2018.
- REBOUÇAS F., P. P.; REBOUÇAS, E. de S.; **MARINHO, L. B.**; SARMENTO, R. M.; TAVARES, J. M. R.S.; DE ALBUQUERQUE, V. H. C. Analysis of human tissue densities:

A new approach to extract features from medical images. *Pattern Recognition Letters*, v. 94, p. 211-218, 2017.

- REBOUÇAS F., P. P.; LIMA M., F. D.; **MARINHO, L. B.**; ALMEIDA, J. S. Novel and powerful mosaic constructor for territorial analysis using mobile robots via Binary Robust Invariant Scalable Keypoints. *Measurement*, v. 114, p. 448-470, 2017.

### **1.5.2 Trabalhos completos publicados em anais de congressos**

- FERREIRA J., M. A. A.; VALENTIM, M.; **MARINHO, L. B.**; NASCIMENTO, N. DE M.M.; SILVA, S. P. P.; ALVES, S. S. A.; RAMALHO, G. L. B.; REBOUCAS F., P. P. Evaluation of Heart Disease Diagnosis Approach using ECG Images. In: *International Joint Conference on Neural Networks (IJCNN)*, 2019, Budapest.
- NASCIMENTO, N. DE M. M.; **MARINHO, L. B.**; PEIXOTO, S. A.; REBOUCAS F., P. P. Heart Arrhythmia Classification Based on Statistical Moments and Structural Co-occurrence. In: *International Joint Conference on Neural Networks (IJCNN)*, 2019, Budapest.
- REBOUCAS F., P.P.; SOUZA, J. W. M.; **MARINHO, L. B.**; HOLANDA, G. B.; S., HARRISON H.; LEITE, A. A. F.; BANDEIRA, T. S.; RODRIGUES, A. W. O. Mobile System to Aid in the Identification and Classification of Electrical Assets using Convolutional Neural Network. In: *2019 8th Brazilian Conference on Intelligent Systems (BRACIS)*, 2019, Salvador.
- DA SILVA, S. P. P.; DA NOBREGA, R. V. M; MEDEIROS, A. G.; **MARINHO, L. B.**; ALMEIDA, J. S.; REBOUCAS F., P. P. Localization of Mobile Robots with Topological Maps and Classification with Reject Option using Convolutional Neural Networks in Omnidirectional Images. In: *2018 International Joint Conference on Neural Networks (IJCNN)*, 2018, Rio de Janeiro.
- DA SILVA, S. P. P.; **MARINHO, L. B.**; ALMEIDA, J. S.; REBOUCAS F., P.P. Robot Localization Using Classification with Reject Option from Structural Co-occurrence Matrix in Omnidirectional Images. In: *2018 7th Brazilian Conference on Intelligent Systems (BRACIS)*, 2018, Sao Paulo.
- HONORIO F., P.; DA SILVA, S. P. P.; ALMEIDA, J. S.; **MARINHO, L. B.**; CARNEIRO, T.; RODRIGUES, A. W. O.; REBOUCAS F., P. P. An Approach to Navigation in Outdoor and Indoor Environments with Unmanned Aerial Vehicle Using Visual Topological Map.

- In: 31st Conference on Graphics, Patterns and Images (SIBGRAPI), 2018, Foz do Iguaçu.
- DA SILVA, S. P. P.; **MARINHO, L. B.**; ALMEIDA, J. S.; REBOUCAS F., P. P. A Novel Approach for Mobile Robot Localization in Topological Maps using Classification with Reject Option from Structural Co-occurrence Matrix. In: 17th International Conference on Computer Analysis of Images and Patterns - CAIP 2017, 2017, Ystad.
  - DA SILVA, S. P. P.; ALMEIDA, J. S.; **MARINHO, L. B.**; REBOUCAS F., P. P. A Novel Approach for Robot Localization in Topological Maps using Classification with Reject Option from Structural Co-occurrence Matrix in Omnidirectional Images. In: 16th IEEE International Conference On Machine Learning And Applications, 2017, Cancun, Mexico.
  - DA SILVA, S. P. P.; ALMEIDA, J. S.; **MARINHO, L. B.**; REBOUCAS F., P. P. A Novel Approach for Robot Localization in Topological Maps using Classification with Reject Option from Structural Co-occurrence Matrix in Omnidirectional Images. In: ICMLA2017, 2017, Cancun.

### ***1.5.3 Capítulos de livros publicados***

- **MARINHO, L. B.**; DE SOUZA J., A. H.; REBOUÇAS F., P. P. A New Approach to Human Activity Recognition Using Machine Learning Techniques. Advances in Intelligent Systems and Computing. 52ed.: Springer International Publishing, 2017, v. , p. 529-538.
- RODRIGUES, M. B.; **MARINHO, L. B.**; NÓBREGA, R. V. M.; SOUZA, J. W. M.; REBOUCAS F., P. P. Lung Segmentation in Chest Computerized Tomography Images Using the Border Following Algorithm. Advances in Intelligent Systems and Computing. 53ed.: Springer International Publishing, 2017, v. , p. 539-548.
- DA SILVA, S. P. P.; **MARINHO, L. B.**; ALMEIDA, J. S.; REBOUÇAS F., P. P. A Novel Approach for Mobile Robot Localization in Topological Maps Using Classification with Reject Option from Structural Co-occurrence Matrix. Lecture Notes in Computer Science. 17ed.: Springer International Publishing, 2017, v. , p. 3-15.

### ***1.5.4 Artigos em processo de avaliação***

- **MARINHO, L. B.**; NASCIMENTO, N. DE M.M.; REBOUÇAS F., P. P.; DE ALBUQUERQUE, V. H. C. Intelligent Edge Computing for Heartbeats Classification and Segmentation using Electrocardiogram Signals: A Novel Deep Learning Approach. Expert Systems with Applications, 2020.

- **MARINHO, L. B.**; NASCIMENTO, N. DE M.M.; RODRIGUES, C. J.; CAVALCANTE, C. H. L.; REBOUÇAS F., P. P.; Sudden Death Classification Using ECG Signals and Structural Co-occurrence Matrix in Patients with Chagas Disease. 34th IEEE International Symposium on Computer-Based Medical Systems, 2021.

### **1.5.5 Registros de softwares**

- **MARINHO, L. B.**; MACEDO, C. C.; HOLANDA, G. B.; MACEDO, H. B.; SOUSA, J. W. M.; REBOUCAS F., P. P.; MACEDO, V. C. Monitoramento e Permanência de pessoas em Eclusas e Elevadores, 2020. Número do registro: BR5120200021-5, data de registro: 27/01/2020, INPI - Instituto Nacional da Propriedade Industrial.
- **MARINHO, L. B.**; MACEDO, C. C.; HOLANDA, G. B.; MACEDO, H. B.; SOUSA, J. W. M.; REBOUCAS F., P. P.; MACEDO, V. C. Detecção de múltiplas massas ao abrir o portão, 2020. Número do registro: BR512020002030-5, data de registro: 27/01/2020. INPI - Instituto Nacional da Propriedade Industrial.
- **MARINHO, L. B.**; MACEDO, C. C.; HOLANDA, G. B.; MACEDO, H. B.; SOUSA, J. W. M.; REBOUCAS F., P. P.; MACEDO, V. C. Desenvolver ferramentas de busca, 2020. Número do registro: BR512020002032-1, data de registro: 27/01/2020. INPI - Instituto Nacional da Propriedade Industrial.
- **MARINHO, L. B.**; MACEDO, C. C.; HOLANDA, G. B.; MACEDO, H. B.; SOUSA, J. W. M.; REBOUCAS F., P. P.; MACEDO, V. C. Controle de acesso temporário de Pedestre por QR-Code utilizando Visão Computacional, 2020. Número do registro: BR512020002031-3, data de registro: 27/01/2020. INPI - Instituto Nacional da Propriedade Industrial.

## **1.6 Organização da tese**

Esta tese está organizada em 5 capítulos. Uma revisão bibliográfica sobre a fisiologia do coração, como o eletrocardiograma é utilizado na identificação da anomalias cardíacas e sobre as Redes Neurais Convolucionais, que são a base teórica do método proposto, são apresentados no Capítulo 2. Posteriormente, o Capítulo 3 aborda a metodologia adotada no estudo das técnicas que usam abordagem clássica para classificação de arritmias cardíacas em sinais de ECG, assim como a avaliação do método proposto na tarefa de segmentação, detecção e classificação dos batimentos cardíacos. No Capítulo 4 são mostrados os resultados obtidos e realizadas as suas

discussões. Por fim, no Capítulo 5, são apresentadas as conclusões e contribuições desta tese.

## 2 FUNDAMENTOS TEÓRICOS

Este capítulo inicia com a elucidação da fisiologia do coração e como o eletrocardiograma é utilizado na identificação das anomalias cardíacas. Em seguida, é feita uma revisão das redes neurais convolucionais, constituindo o arcabouço teórico para a compreensão do método propostos e para a avaliação dos resultados obtidos.

### 2.1 Aspectos do coração e eletrocardiograma

A eletrocardiografia é um procedimento não invasivo usado em diagnósticos clínicos do coração. Seu registro, denominado ECG, mostra a série de ondas relacionadas aos impulsos elétricos que ocorrem durante cada batimento desse músculo. Os resultados podem ser impressos em papel ou exibidos em um monitor (ZIPES, 2012).

A contração do músculo cardíaco é o resultado direto da excitação elétrica celular descrita pelo ECG. A despolarização inicia a contração de cada célula muscular individual. Dessa forma, a ativação elétrica de cada célula é uma indicação do seu funcionamento. Portanto, o ECG é o resultado da despolarização do músculo cardíaco de forma repetitiva controlada. Ao rastrear o processo de despolarização elétrica das células do músculo cardíaco, uma impressão da funcionalidade do coração pode ser formada e usada para reconhecer regiões na estrutura do coração que não estão funcionando de acordo com as especificações e podem exigir atenção médica. Qualquer desvio do ECG típico observado no sinal de despolarização elétrica registrado é analisado e classificado como um determinado distúrbio cardíaco (NAJARIAN; SPLINTER, 2012).

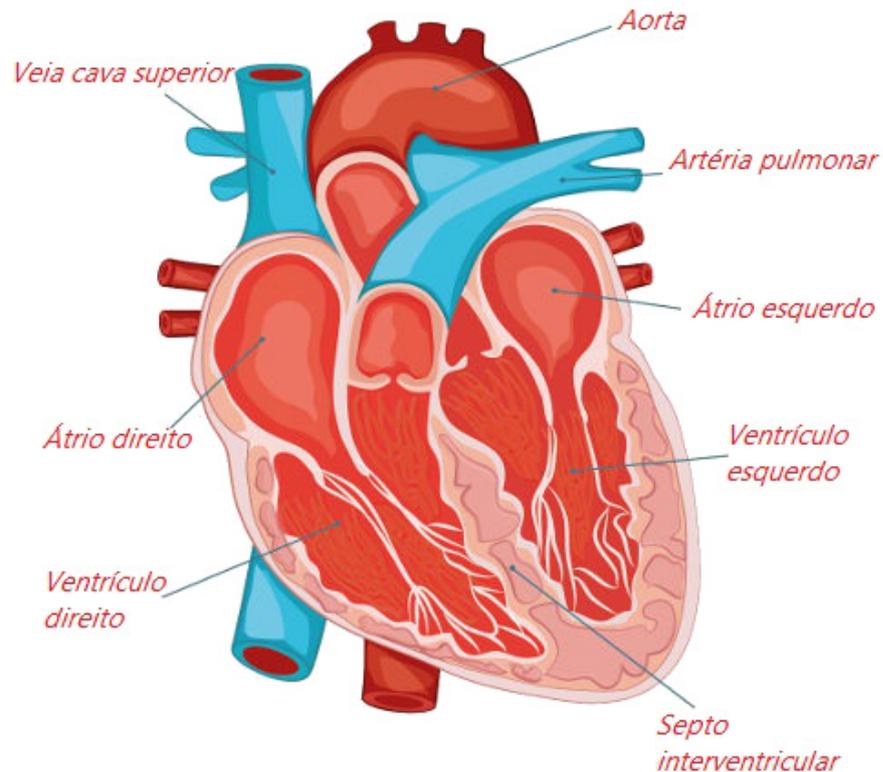
#### 2.1.1 Funções e estrutura do coração

O coração é formado por músculos responsáveis pela circulação do sangue pelo corpo. Sua anatomia e o sistema de condução são descritos na Figura 2. Ele tem quatro funções principais: coletar o sangue que precisa ser refinado de todas as partes do corpo (através das veias), bombear o sangue coletado para os pulmões, coletar o sangue refinado dos pulmões e bombear o sangue refinado de volta para todas as partes do corpo (BUCKBERG *et al.*, 2018).

Como pode ser visto na Figura 2, o coração tem quatro câmaras: dois átrios e dois

<sup>1</sup> Disponível em: <<https://courses.lumenlearning.com/suny-contemporaryhealthissues/chapter/internal-structure-of-the-heart/>>. Acesso em 21 de setembro de 2020.

Figura 2 – Anatomia e sistema de condução do coração.



Fonte: Adaptado de *Cardiovascular System Module 3: Heart Anatomy*.<sup>1</sup>

ventrículos. O átrio é separado do sistema venoso por uma válvula para que o fluxo só seja possível em uma direção. A veia cava superior e a veia cava inferior conduzem ao átrio direito em combinação com o seio coronário, enquanto as veias pulmonares irrigam o átrio esquerdo. Quando o átrio se contrai, ele bombeia o sangue retido para o ventrículo, que também é separado por uma válvula. A válvula permite apenas o fluxo do átrio para o ventrículo e não na direção oposta. Essa válvula é chamada de válvula atrioventricular (WELLENS, 2012).

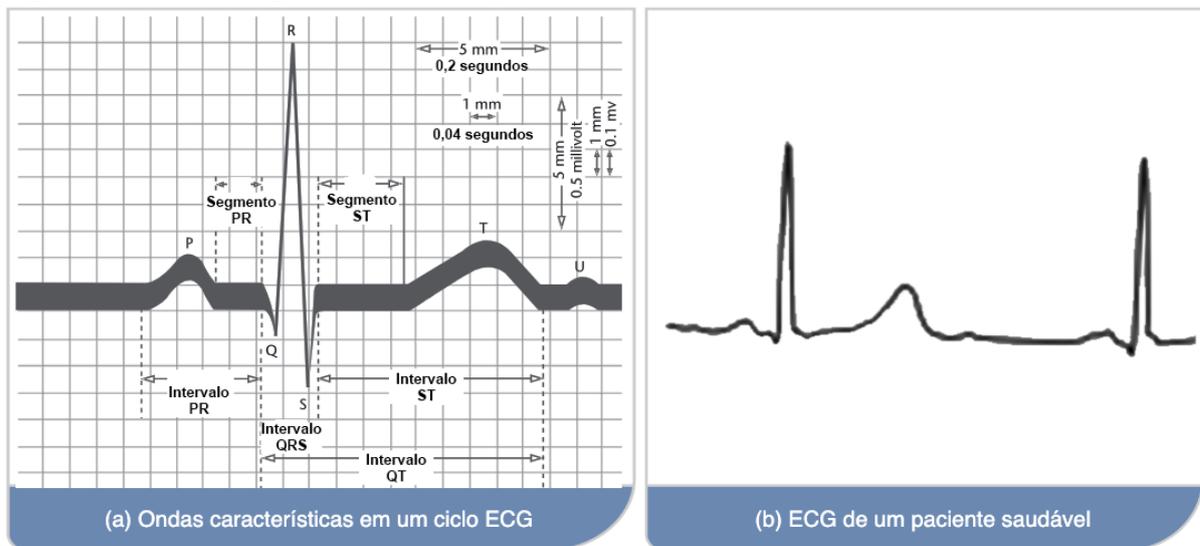
O ventrículo se expande quando é preenchido com a pressão gerada pelo fluxo sanguíneo do átrio. Ele se contrai fora de sincronia com o átrio como resultado de uma linha de retardo elétrico entre a excitação do átrio e o ventrículo. O ventrículo bombeia o sangue para uma artéria, que novamente é separada do ventrículo por uma válvula unidirecional. O ventrículo esquerdo bombeia sangue para a aorta, que é separada do ventrículo pela válvula aórtica para evitar o refluxo. A aorta fornece o sangue refinado ao corpo. O ventrículo direito espreme o sangue que precisa ser refinado para a artéria pulmonalis, separada pela válvula pulmonar (NAJARIAN; SPLINTER, 2012).

### 2.1.2 Sinal eletrocardiograma

O processo de despolarização e repolarização durante cada ciclo cardíaco gera diferenças de potencial elétrico local, que podem ser medidas na pele por meio de equipamento de registro eletrônico. Este grupo de sinais, denominado ECG, constitui o sinal clínico mais informativo comumente utilizado no diagnóstico do sistema cardiovascular (NAJARIAN; SPLINTER, 2012).

O ECG representa o padrão repetitivo de despolarização e repolarização elétrica do músculo cardíaco. A forma esquemática de um período do ECG é ilustrada na Figura 3(a), e um sinal típico de ECG saudável é mostrado na Figura 3(b). O ECG é caracterizado por cinco picos, representados pelas letras P, Q, R, S, T, e às vezes seguido por um sexto pico, a onda U. A onda P é o resultado da despolarização do átrio, enquanto as demais ondas são causadas pelo ventrículo (CLIFFORD *et al.*, 2006).

Figura 3 – Forma esquemática de um período do ECG e amostra de ECG saudável.



Fonte: Adaptado de Clifford *et al.* (2006).

A eletrocardiografia foi introduzida por meio dos esforços pioneiros do cientista holandês Willem Einthoven em 1903. Ele usou um galvanômetro para projetar uma maneira de registrar os potenciais de ação. Ele também introduziu os marcadores P, Q, R, S e T no ECG padrão. Os ECGs iniciais foram registrados diretamente em papel e, de fato, ainda estão em muitos laboratórios de eletrofisiologia cardíaca clínica. O galvanômetro foi acoplado diretamente a uma caneta de tinta. Dessa forma, uma tensão que leva a uma deflexão do galvanômetro moveria ou direcionaria a caneta sobre o papel. Cada eletrodo individual tinha seu próprio

galvanômetro e caneta de tinta separada. Esse método ainda é o padrão ouro para gravações analógicas. Porém, atualmente, os eletrodos são conectados a amplificadores e filtros (PASTORE *et al.*, 2016).

A despolarização do átrio produz a onda P, enquanto a despolarização do ventrículo gera o complexo QRS. Além disso, a repolarização ventricular cria a onda T. O significado da onda U é incerto, mas pode ser por conta da repolarização do sistema de Purkinje. O intervalo PR não deve passar de 0,20 segundos no papel ECG, onde cada quadrado pequeno representa 0,04 segundos. Já o limite superior da duração normal do complexo QRS é no máximo 0,12 segundos. Uma duração menor que 0,12 segundos significa que o impulso foi iniciado acima do nó A-V (supraventricular) ou acima do mesmo. Uma duração de QRS maior que 0,12 segundos pode indicar que o impulso proveniente do ventrículo foi originado do tecido supraventricular, mas com condução prolongada através do ventrículo. Ademais, a onda T é a recuperação ventricular ou repolarização dos ventrículos (MADEIRO *et al.*, 2018).

A interpretação das arritmias cardíacas consiste na análise da duração e configuração entre as ondas P, intervalo PR e o complexo QRS. O ECG deve ser analisado quanto a frequência, ritmo, locus do marcapasso dominante e as configuração da onda P e do complexo QRS (MADEIRO *et al.*, 2018).

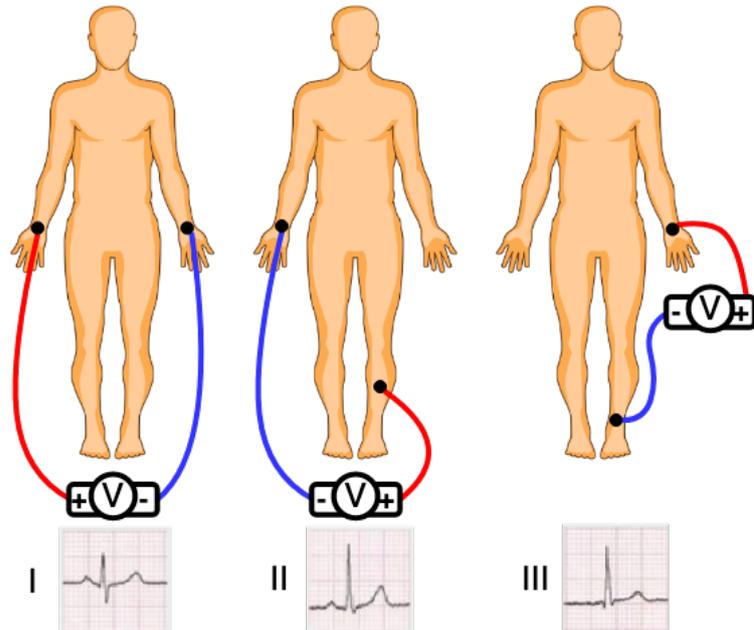
### **2.1.3 Disposição dos eletrodos para o eletrocardiograma**

O ECG não é medido diretamente no coração, mas no exterior do corpo. Existem várias disposições de eletrodos internacionalmente aceitas, que são descritas a seguir.

O padrão mais antigo de medição de ECG é um registro de extremidades de três pontos, nos braços e em uma perna, proposta por Einthoven (EITHOVEN, 1903). O registro usa três medidas bipolares que são as seguintes: o registro I está entre os braços esquerdo (L, *left*) e direito (R, *right*), ou seja,  $V_L - V_R$ ; a gravação II ocorre entre a perna esquerda (F, *foot*) e o braço direito, ou seja,  $V_F - V_R$ ; e a gravação III está entre a perna esquerda e o braço esquerdo, ou seja,  $V_F - V_L$ . Esta configuração de ECG padrão para a colocação do eletrodo é ilustrada na Figura 4.

Outra distribuição popular dos eletrodos é descrito na Figura 5. Uma curva de deslocamento no tórax delinea a colocação do eletrodo começando da direita, com um eletrodo de cada lado do esterno e mais três eletrodos seguindo ao longo do lado esquerdo do esterno na quinta costela com o sexto eletrodo sob a axila na sexta costela. Esta colocação de eletrodo usando seis eletrodos é chamada de V1 – V6. Esses seis eletrodos são adicionados aos das três

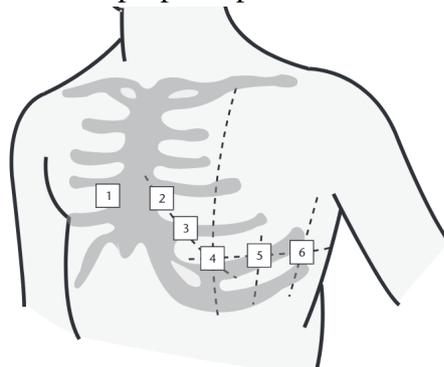
Figura 4 – Disposição tradicional dos eletrodos proposta por Einthoven.



Fonte: elaborado pelo autor.

extremidades. Este padrão de ECG é chamado de colocação de Wilson (WILSON *et al.*, 1934). Como pode ser visto, nesta norma, a colocação dos eletrodos segue o contorno do coração de cima para baixo.

Figura 5 – Disposição do eletrodo torácico proposta por Wilson.

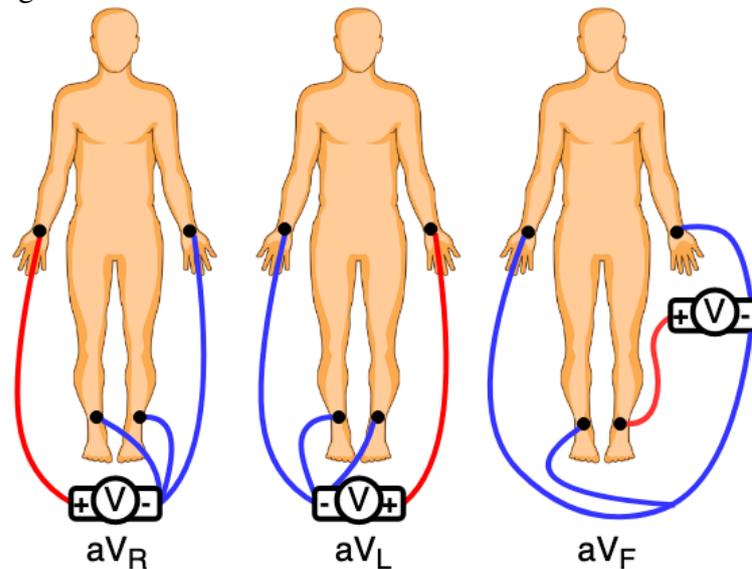


Fonte: (CLIFFORD *et al.*, 2006)

Posteriormente, Goldberger descobriu que a sensibilidade das medições pode ser melhorada desacoplando a extremidade na qual o sinal é desejado e fazendo a medição diferencial com outros eletrodos (GOLDBERGER, 2018). A Figura 6 mostra esta disposição.

A combinação de todas as medições anteriores dá um total de 12 gravações: três eletrodos nas extremidades (Einthoven), seis no peito (Wilson) e as três gravações desacopladas introduzidas por Goldberger. Essas são as gravações tipicamente obtidas durante um exame

Figura 6 – Disposição dos eletrodos proposta por Goldberger.



Fonte: elaborado pelo autor.

cardíaco de rotina.

#### 2.1.4 Doenças cardiovasculares e eletrocardiograma

Os desvios da funcionalidade normal do sistema cardiovascular estão associados a certas condições patológicas, que podem ser genéticas ou devido a disfunções, como infecções, falta de oxigênio e obstrução dos vasos sanguíneos que fornecem sangue ao próprio coração. A seguir, algumas das principais doenças cardiovasculares são brevemente introduzidas e as mudanças no ECG e doenças relacionadas são discutidas. Conhecer as doenças cardiovasculares permitir desenvolver métodos computacionais para o processamento de ECG que podem detectar anormalidades em seus estágios iniciais.

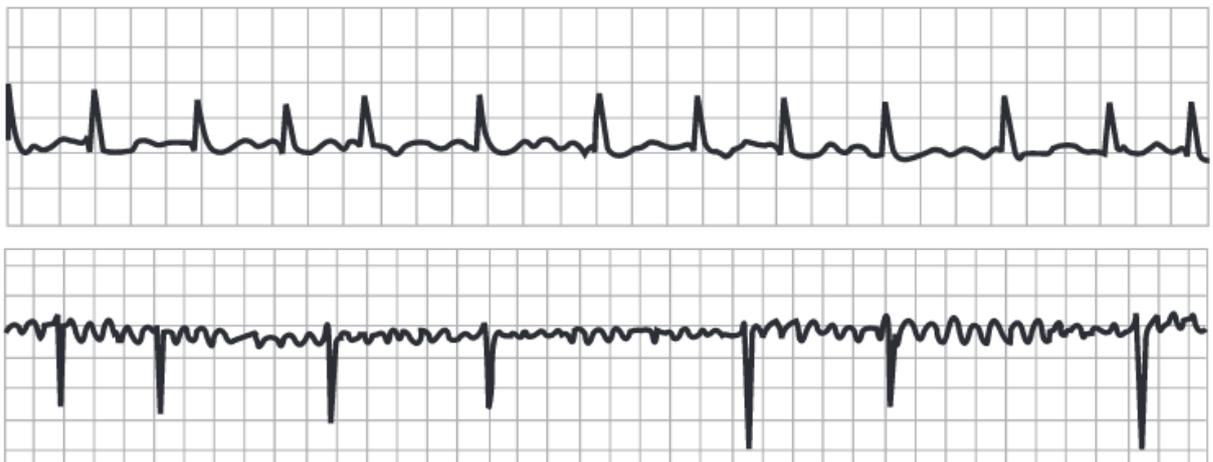
#### *Fibrilação Atrial*

A Fibrilação Atrial (FA) é uma das arritmias mais comuns que ocorrem como resultado de doença reumática, infecções, doença aterosclerótica ou hipertireoidismo. Essas condições causadas pela FA não são tão fatais quanto algumas arritmias ventriculares, mas fornecem um risco aumentado de acidente vascular cerebral. Os sintomas físicos incluem palpitações, tonturas e falta de ar. Algumas pessoas com FA nunca notam qualquer sinal de desconforto (REESE-PETERSEN *et al.*, 2020).

A FA é caracterizada pela ativação elétrica atrial descoordenada, em taxas entre

350 a 600 descargas/min, com conseqüente deterioração mecânica, fibrose associada e perda de massa muscular atrial, afetando a capacidade dos átrios de bombear sangue de forma eficaz. Uma vez que as fibras musculares atriais se contraem independentemente, não há ondas P no ECG, apenas ondas fibrilares irregulares (onda f), que podem contribuir para a formação de coágulos sanguíneos e eventos tromboembólicos, aumentando em cinco vezes o risco de acidente vascular cerebral. As ondas f têm características de amplitudes baixas e variáveis, e ondulações irregulares na linha de base. A atividade contínua da frente de onda que atinge o nó AV promove batimentos ventriculares irregulares e, conseqüentemente, contrações irregulares. A frequência ventricular média é superior a 100pbm. Como o sistema de condução ventricular é normal, o complexo QRS e a onda T têm forma normal. No entanto, a detecção do final da onda T é um desafio na AF, uma vez que a onda T e a onda f são motivo de confusão (MADEIRO *et al.*, 2018). A Figura 7 mostra um exemplo de um ECG com AF.

Figura 7 – Exemplo de ECG com Fibrilação Atrial.



Fonte: (CLIFFORD *et al.*, 2006)

### ***Arritmia Ventricular***

Nas arritmias ventriculares, a ativação ventricular não se origina no nó A-V. Além disso, a arritmia não prossegue nos ventrículos de forma normal. Em um coração normal, no qual a ativação prossegue para os ventrículos ao longo do sistema de condução, as paredes internas dos ventrículos são ativadas quase simultaneamente, e a frente de ativação prossegue principalmente radialmente em direção às paredes externas. Como resultado, o complexo QRS de um coração saudável é de duração relativamente curta (NAJARIAN; SPLINTER, 2012).

Nas arritmias ventriculares, entretanto, como o sistema de condução ventricular está

rompido ou a ativação ventricular começa longe do nó A-V, leva mais tempo para que a frente de ativação prossiga por todo o músculo ventricular. Isso resulta em um complexo QRS mais longo. O critério para ativação ventricular normal requer que o intervalo QRS seja menor que 0,1s. Um intervalo QRS com duração superior a 0,1s indica ativação ventricular anormal. Um exemplo de arritmia ventricular é ilustrado na Figura 8 (GOLDBERGER, 2018).

Figura 8 – ECG de um paciente com arritmia ventricular.



Fonte: (GOLDBERGER, 2018)

Outra característica do distúrbio ventricular é a contração ventricular prematura. Uma contração ventricular prematura é aquela que ocorre anormalmente cedo. Se a origem do distúrbio for no músculo ventricular, o complexo QRS tem uma forma muito anormal e dura mais do que 0,1 s. Normalmente, a onda P não está associada a ele. O complexo arritmogênico produzido por essa arritmia supraventricular dura menos de 0,1s (NAJARIAN; SPLINTER, 2012).

### ***Taquicardia Ventricular***

Um ritmo de origem ventricular pode ser consequência de uma condução mais lenta no músculo ventricular isquêmico que leva à ativação circular. Isso resulta na ativação dos músculos ventriculares em uma taxa alta (mais de 120 batimentos por minuto), causando complexos QRS rápidos e amplos. Essa arritmia é chamada de Taquicardia Ventricular (TV).

Essa anomalia geralmente é uma consequência de isquemia e infarto do miocárdio. A principal alteração no ECG que indica a ocorrência de VT é a frequência cardíaca muito rápida que pode ser detectada no domínio de Fourier (GOLDBERGER, 2018).

### ***Fibrilação Ventricular***

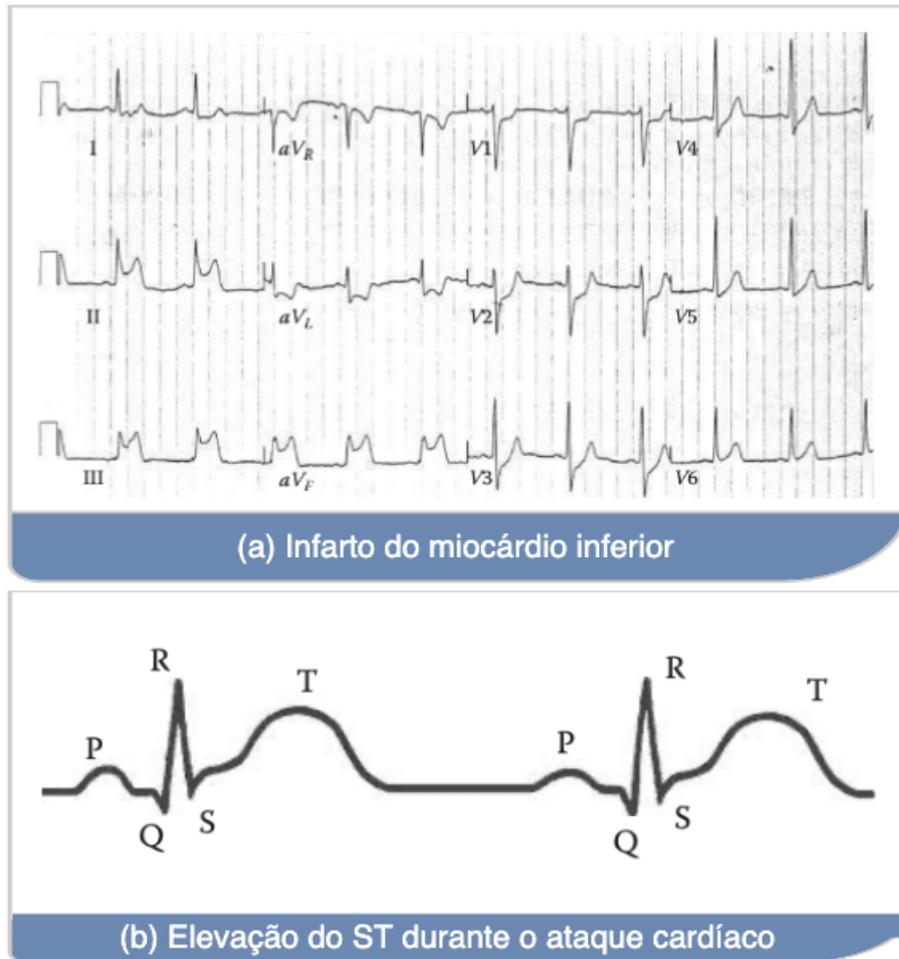
Quando a despolarização ventricular ocorre de forma caótica, a situação é chamada de fibrilação ventricular. Isso se reflete no ECG, que demonstra ondulações irregulares grosseiras sem complexo QRS. A causa da fibrilação é o estabelecimento de várias alças de reentrada, geralmente envolvendo o músculo cardíaco doente. Nesse tipo de arritmia, a contração do músculo ventricular também é irregular e, portanto, o momento é ineficaz no bombeamento do sangue. A falta de circulação sanguínea leva à perda quase imediata da consciência e até à morte em minutos. A fibrilação ventricular pode ser interrompida com pulso de desfibrilador externo e medicação apropriada (NAJARIAN; SPLINTER, 2012).

### ***Infarto do Miocárdio***

Se uma artéria coronária é obstruída, o transporte de oxigênio para o músculo cardíaco é diminuído, causando um débito de oxigênio no músculo, que é chamado de isquemia. A isquemia causa alterações no potencial de repouso e na repolarização das células musculares. Essa anormalidade é observada no ECG como mudanças na forma da onda T. Se o transporte de oxigênio for encerrado em uma determinada área, o músculo cardíaco morre nessa região. Isso é chamado de infarto do miocárdio ou ataque cardíaco. Após um bloqueio nos vasos sanguíneos que fornecem oxigênio e nutrientes ao músculo cardíaco, as células musculares da região ficam gravemente comprometidas. Algumas células podem morrer enquanto outras sofrerão danos graves, resultando em uma diminuição da capacidade de conduzir impulsos, gerando sua própria despolarização. As células mortas serão eventualmente substituídas por colágeno, pois o coração não tem a capacidade de se regenerar (CLIFFORD *et al.*, 2006).

A Figura 9 (a) mostra gravações de um ECG que apresenta o infarto do miocárdio. Em muitos casos de ataque cardíaco, devido à existência de células morrendo no músculo cardíaco, não haverá queda significativa entre o complexo QRS e a onda T. O período entre a parte S e a onda T também parecerá contínuo. Isso é conhecido como elevação ST. A Figura 9 (b) ilustra o efeito das células mortas no potencial ST. A elevação do segmento ST é um dos indicadores mais reconhecíveis de infarto do miocárdio.

Figura 9 – ECG de um paciente com infarto do miocárdio inferior (a) e ilustração esquemática da elevação do ST durante o ataque cardíaco (b).



Fonte: (NAJARIAN; SPLINTER, 2012)

### ***Flutter Atrial***

Quando a frequência cardíaca é suficientemente elevada para que o intervalo iso-elétrico entre o final de T e o início de P desapareça, a arritmia é chamada de Flutter Atrial. Acredita-se também que a origem envolve uma via atrial de reentrada. A frequência dessas flutuações está entre 220 e 300 batimentos por minuto. O nó A-V e depois os ventrículos são geralmente ativados a cada segundo ou a cada terceiro impulso atrial (NAJARIAN; SPLINTER, 2012).

## 2.2 Introdução ao Aprendizado Profundo

O aprendizado profundo (*Deep Learning*) é formado por um conjunto de métodos que modelam dados com arquiteturas complexas combinando diferentes transformações não lineares. A base desse aprendizado são as Redes Neurais Artificiais (*Artificial Neural Networks*, ANNs), que são combinadas para formar as redes neurais profundas (GOODFELLOW *et al.*, 2016; BISHOP, 2006).

Essas técnicas permitiram um progresso significativo nas áreas de processamento de sinais e imagens, incluindo reconhecimento facial, reconhecimento de fala e processamento de linguagem natural. As aplicações potenciais são numerosas. Um exemplo notável é o programa AlphaGo, que aprendeu a jogar *Go* pelo método de aprendizado profundo e venceu o campeão mundial em 2016 (LAPAN, 2018).

Há vários tipos de arquiteturas de redes neurais artificiais que se baseiam em uma cascata profunda de camadas: Perceptron Multicamadas (*Multilayer Perceptrons*, MLPs), Redes Neurais Convolucionais (*Convolutional Neural Networks*, CNNs) e Redes Neurais Recorrentes (*Recurrent Neural Networks*, RNNs). Na prática, os modelos de Aprendizado Profundo trabalham com algoritmos de otimização estocástica para evitar ficar preso em mínimos locais, variar a superfície de erro e acelerar o processo de convergência trabalhando com *batch* (LAPAN, 2018).

## 2.3 Redes Neurais Artificiais

As Redes Neurais Artificiais (*Artificial Neural Networks*, ANNs) baseiam-se em uma coleção de unidades conectadas ou nós chamados neurônios artificiais, que modelam vagamente os neurônios em um cérebro biológico. Cada conexão, como as sinapses em um cérebro biológico, pode transmitir um sinal a outros neurônios. Um neurônio artificial que recebe um sinal, o processa e pode sinalizar os neurônios conectados a ele. O sinal em uma conexão é um número real, e a saída de cada neurônio é calculada por alguma função não linear da soma de suas entradas,  $y = f(\mathbf{x}, \theta)$  (BRAGA *et al.*, 2007).

Como costuma ocorrer na aprendizado estatístico, os parâmetros  $\theta$  são estimados a partir de uma amostra dos dados. A função de minimização não é convexa, levando a minimizadores locais. O sucesso do método veio de um teorema de aproximação universal devido a Cybenko (1989) e Hornik (1991). Além disso, LeCun *et al.* (1988) propôs uma maneira eficiente de calcular o gradiente de uma rede neural, chamada *backpropagation* do gradiente,

que permite obter facilmente um minimizador local do critério quadrático.

### 2.3.1 Neurônio Artificial

Um neurônio artificial é uma função  $f_j$  de entrada  $\mathbf{x} = (x_1, \dots, x_d)$  ponderada por um vetor de pesos de conexão  $\mathbf{w}_j = (w_{j,1}, \dots, w_{j,d})$ , complementada por um viés (*bias*) de neurônio  $b_j$ , e associado a uma função de ativação  $\phi$ ,

$$y_j = f_j(\mathbf{w}) = \phi(\langle \mathbf{w}_j, \mathbf{x} \rangle + b_j). \quad (2.1)$$

Várias funções de ativação podem ser consideradas. A Tabela 1 mostra as mais utilizadas. Na Figura 10, a representação esquemática de um neurônio artificial é apresentada em (a) e os gráficos das funções de ativação em (b).

Tabela 1 – Exemplos de funções de ativação

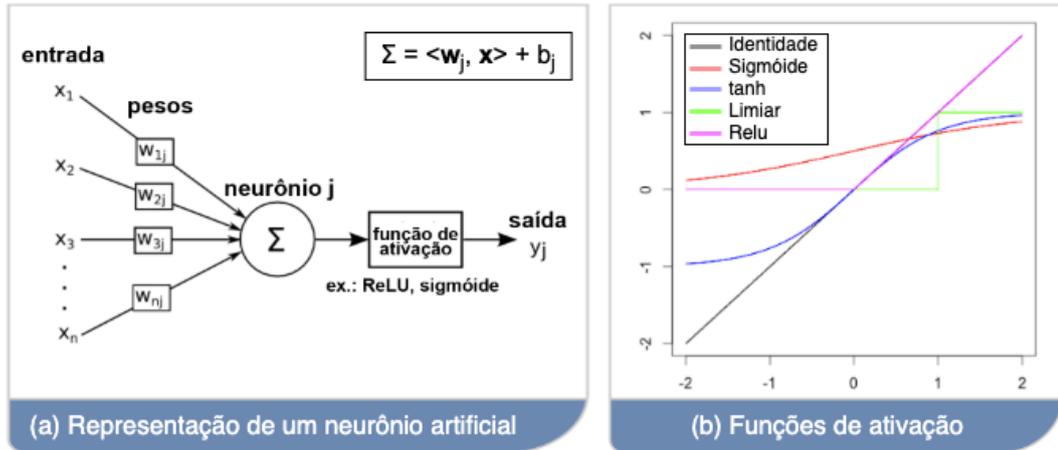
Função	Definição
<b>Identidade</b>	$\phi(\mathbf{x}) = \mathbf{x}$ ,
<b>Sigmóide (ou logística)</b>	$\phi(\mathbf{x}) = \frac{1}{1 + \exp(-\mathbf{x})}$ ,
<b>Tangente hiperbólica (tanh)</b>	$\phi(\mathbf{x}) = \frac{\exp(\mathbf{x}) - \exp(-\mathbf{x})}{\exp(\mathbf{x}) + \exp(-\mathbf{x})} = \frac{\exp(2\mathbf{x}) - 1}{\exp(2\mathbf{x}) + 1}$ ,
<b>Limiar rígido</b>	$\phi_\beta(\mathbf{x}) = \mathbb{1}_{\mathbf{x} \geq \beta}$ ,
<b>Unidade Linear Retificada (ReLU)</b>	$\phi(\mathbf{x}) = \max(0, \mathbf{x})$ .

Fonte: elaborado pelo autor.

Historicamente, a sigmóide foi a função de ativação mais utilizada, pois é diferenciável e permite manter os valores no intervalo  $[0, 1]$ . No entanto, seu gradiente é muito próximo de 0 quando  $|\mathbf{x}|$  não está próximo a 0. A Figura 11 mostra a função sigmóide e sua derivada.

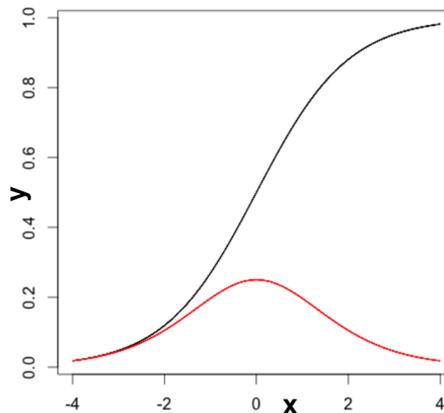
No caso do aprendizado profundo que é uma rede neural com um grande número de camadas, isso causa problemas para o algoritmo de treinamento estimar os parâmetros. Por conta disto, a função sigmóide foi substituída pela função de Unidade Linear Retificada. A função ReLU não é diferenciável em 0, mas na prática isso não é realmente um problema, pois a probabilidade de ter uma entrada igual a 0 é geralmente nula. Esse problema costuma ser contornado adicionando uma pequena polarização positiva para garantir que cada unidade esteja ativa. Diversas variações da função ReLU são considerados para certificar de que todas as

Figura 10 – Representação esquemática de um neurônio artificial e gráficos de funções de ativação.



Fonte: Adaptado de (BRAGA *et al.*, 2007).

Figura 11 – Função sigmóide (preto) e sua derivada (vermelho).



Fonte: elaborado pelo autor.

unidades tenham um gradiente que não desaparece e que quando  $x < 0$  a derivada não seja igual a 0. Isto é,

$$\phi(\mathbf{x}) = \max(\mathbf{x}, 0) + \alpha \min(\mathbf{x}, 0), \quad (2.2)$$

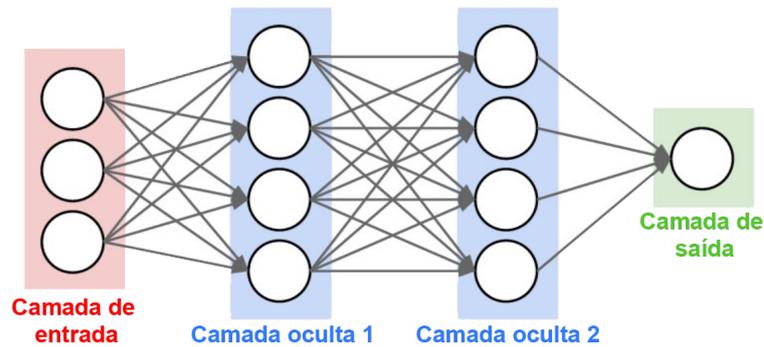
onde  $\alpha$  é um parâmetro fixo definido com um pequeno valor positivo ou um parâmetro a ser estimado.

### 2.3.2 Perceptron Multicamadas

O Perceptron Multicamadas (*Multilayer Perceptrons*, MLPs) é uma estrutura composta por várias camadas ocultas de neurônios onde a saída de um neurônio de uma camada torna-se a entrada de um neurônio da próxima camada. Além disso, a saída de um neurônio

também pode ser a entrada de um neurônio da mesma camada ou de um neurônio de camadas anteriores (como ocorre nas redes neurais recorrentes). Na última camada, chamada de camada de saída, pode-se aplicar uma função de ativação diferente das camadas ocultas dependendo do tipo de problema: regressão ou classificação. A Figura 12 apresenta a arquitetura de uma MLP com três neurônios de entrada, um neurônio de saída e duas camadas ocultas (BRAGA *et al.*, 2007; DUDA *et al.*, 2012).

Figura 12 – Arquitetura de um Perceptron Multicamadas.



Fonte: elaborado pelo autor.

A MLP têm uma arquitetura básica, pois cada neurônio de uma camada está ligada a todos os da próxima camada, mas não tem nenhuma ligação com os neurônios da mesma camada. Os parâmetros da arquitetura são o número de camadas ocultas e de neurônios em cada camada. As funções de ativação também são escolhidas pelo usuário. Na camada de saída, a função de ativação é geralmente diferente daquela usada nas camadas ocultas. Em problemas de regressão, nenhuma função de ativação é aplicada na camada de saída. Para classificação binária, a saída fornece uma classificação de  $P(Y = 1/X)$ , uma vez que este valor está em  $[0, 1]$ , a função de ativação sigmóide é geralmente considerada. Para classificação multiclasse, a camada de saída contém um neurônio por classe  $i$ , dando uma classificação de  $P(Y = i/X)$ . A soma de todos esses valores deve ser igual a 1. A função multidimensional softmax é geralmente usada,

$$\text{softmax}(z)_i = \frac{\exp(z_i)}{\sum_j \exp(z_j)}. \quad (2.3)$$

De forma resumida, a formulação matemática de um Perceptron Multicamadas com  $k = 1, \dots, CO$  camadas ocultas, considerando  $h^{(0)}(\mathbf{x}) = \mathbf{x}$  é,

$$\begin{aligned} a^{(k)}(\mathbf{x}) &= b^{(k)} + \mathbf{W}^{(k)} h^{(k-1)}(\mathbf{x}) \\ h^{(k)}(\mathbf{x}) &= \phi(a^{(k)}(\mathbf{x})). \end{aligned} \quad (2.4)$$

Para  $k = CO + 1$  camadas de saída,

$$\begin{aligned} a^{(CO+1)}(\mathbf{x}) &= b^{(CO+1)} + \mathbf{W}^{(CO+1)}h^{(CO)}(\mathbf{x}) \\ h^{(CO+1)}(\mathbf{x}) &= \phi_s(a^{(CO+1)}(\mathbf{x})) := f(\mathbf{x}, \boldsymbol{\theta}), \end{aligned} \tag{2.5}$$

em que  $\phi$  é a função de ativação da camada oculta e  $\phi_s$  é a função de ativação da camada de saída (por exemplo, softmax para classificação multiclasse). Em cada etapa,  $\mathbf{W}^{(k)}$  é uma matriz com o número de linhas igual ao número de neurônios na camada  $k$  e número de colunas igual ao número de neurônios na camada  $k - 1$ .

### ***Estimação dos parâmetros***

Uma vez escolhida a arquitetura da rede, os parâmetros (os pesos  $\mathbf{w}_j$  e o *bias*  $b_j$ ) devem ser estimados a partir de uma amostra de aprendizagem. A estimativa é obtida minimizando uma função de perda com um algoritmo de gradiente descendente. Dessa forma, deve-se começar pela escolha da função de perda.

### ***Funções de perda***

É comum estimar os parâmetros maximizando a probabilidade (ou equivalentemente o logaritmo da probabilidade). Isso corresponde à minimização da função de perda, que é o oposto do log da verossimilhança. Seja  $\boldsymbol{\theta}$  o vetor de parâmetros a estimar, considera-se a função de perda esperada como

$$L(\boldsymbol{\theta}) = -\mathbb{E}_{(X,Y) \sim P}(\log(p_{\boldsymbol{\theta}}(Y/X))). \tag{2.6}$$

Se o modelo é Gaussiano, ou seja,  $p_{\boldsymbol{\theta}}(Y/X = x) \sim \mathcal{N}(f(x, \boldsymbol{\theta}), I)$ , maximizar a probabilidade é equivalente a minimizar a perda quadrática

$$L(\boldsymbol{\theta}) = \mathbb{E}_{(X,Y) \sim P}(\|Y - f(X, \boldsymbol{\theta})\|^2). \tag{2.7}$$

Para classificação binária, com  $Y \in \{0, 1\}$ , maximizar a probabilidade logarítmica corresponde à minimização da entropia cruzada. Para  $f(X, \boldsymbol{\theta}) = p_{\boldsymbol{\theta}}(Y = 1/X)$ ,

$$L(\boldsymbol{\theta}) = \mathbb{E}_{(X,Y) \sim P}[Y \log(f(X, \boldsymbol{\theta})) + (1 - Y) \log(1 - f(X, \boldsymbol{\theta}))]. \tag{2.8}$$

Esta função de perda é bem adaptada com a função de ativação sigmóide, visto que o uso do logaritmo evita valores muito pequenos para o gradiente. Finalmente, para um problema de classificação multiclasse, considera-se uma generalização da função de perda anterior para  $k$  classes

$$L(\theta) = -\mathbb{E}_{(X,Y)\sim P} \left[ \sum_{j=1}^k \mathbb{1}_{Y=j} \log p_{\theta}(Y = j/X) \right]. \quad (2.9)$$

O objetivo é minimizar o erro de classificação, mas o mesmo não é suave. Assim, considera-se a entropia cruzada (ou eventualmente um substituto convexo).

### ***Risco empírico com penalização***

A perda esperada pode ser escrita como,

$$L(\theta) = \mathbb{E}_{(X,Y)\sim P} [\ell(f(X, \theta), Y)], \quad (2.10)$$

e está associado a uma função de perda  $\ell$ . Para estimar os parâmetros  $\theta$ , usa-se uma amostra de treinamento  $(\mathbf{X}_i, \mathbf{Y}_i)_{1 \leq i \leq n}$  e minimiza-se a perda empírica,

$$\tilde{L}_n(\theta) = \frac{1}{n} \sum_{i=1}^n \ell(f(\mathbf{X}_i, \theta), \mathbf{Y}_i), \quad (2.11)$$

normalmente, adiciona-se um termo de regularização. Isso minimiza o risco empírico com uma penalização

$$L_n(\theta) = \frac{1}{n} \sum_{i=1}^n \ell(f(\mathbf{X}_i, \theta), \mathbf{Y}_i) + \lambda \Omega(\theta). \quad (2.12)$$

Pode-se considerar a regularização  $\mathbb{L}^2$ . Usando as mesmas notações da Seção 2.3.2,

$$\begin{aligned} \Omega(\theta) &= \sum_k \sum_i \sum_j (\mathbf{W}_{i,j}^{(k)})^2 \\ &= \sum_k \left\| \mathbf{W}^{(k)} \right\|_F^2, \end{aligned} \quad (2.13)$$

em que  $\|\mathbf{W}^{(k)}\|_F$  denota a norma de Frobenius da matriz  $\mathbf{W}$ . Observa-se que apenas os pesos são penalizados, ou seja, os *biases* não são penalizados. É fácil calcular o gradiente de  $\Omega(\theta)$ :

$$\nabla_{\mathbf{W}^{(k)}} \Omega(\theta) = 2\mathbf{W}^{(k)}. \quad (2.14)$$

A regularização  $\mathbb{L}^1$  também pode ser considerada,

$$\Omega(\theta) = \sum_k \sum_i \sum_j |\mathbf{W}^{(k)}(i,j)|. \quad (2.15)$$

Para minimizar  $L_n(\theta)$ , um algoritmo de gradiente descendente estocástico é usado. Para calcular o gradiente, um método inteligente, denominado algoritmo *Backpropagation*, é considerado (LECUN *et al.*, 1988). Essa técnica é ainda crucial para o aprendizado profundo.

## 2.4 Redes Neurais Convolucionais

As Redes Neurais Convolucionais (*Convolutional Neural Networks*, CNNs) são uma classe de algoritmo de *Deep Learning* projetadas para processar matrizes como imagens. Ao contrário dos algoritmos de visão computacional clássicos, as CNNs podem operar diretamente em dados brutos e com menor pré-processamento. Outra particularidade dessas redes, em relação aos métodos primitivos de visão computacional, é sua capacidade de aprender as características relevantes para cada problema (LECUN *et al.*, 1998).

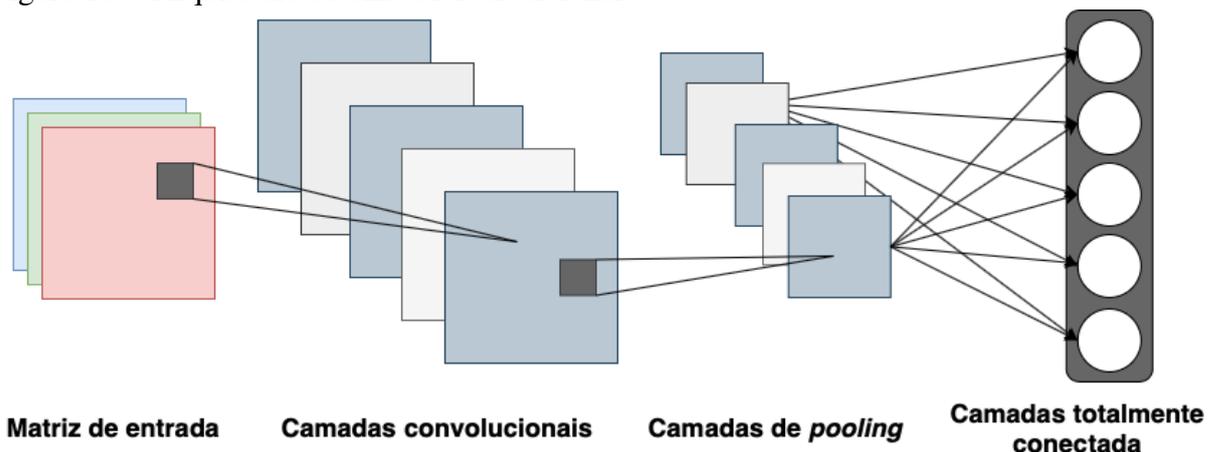
As camadas convolucionais em uma CNN assemelham-se à estrutura do córtex visual humano, onde uma série de camadas processam uma imagem de entrada e identificam características progressivamente mais complexas. Assim, cada uma dessas camadas convolucionais empilhadas são capazes de aprender formas mais sofisticadas (BISHOP, 2006).

As ANNs não se adaptam bem com imagens completas. No conjunto de dados *Canadian Institute For Advanced Research* (CIFAR-10), por exemplo, mesmo com o tamanho das imagens sendo  $32 \times 32 \times 3$  (32 canais de largura, 32 de altura e 3 canais de cor), um único neurônio totalmente conectado em uma primeira camada oculta de uma ANN teria  $32 \times 32 \times 3 = 3072$  pesos. Essa quantidade ainda parece administrável, mas claramente essa estrutura totalmente conectada não se ajusta a imagens maiores. Por outro lado, uma entrada maior, *e.g.*  $200 \times 200 \times 3$ , exigiria neurônios com  $200 \times 200 \times 3 = 120.000$  pesos. Notoriamente, essa conectividade é um desperdício e o grande número de parâmetros levaria rapidamente ao sobreajuste (*overfitting*).

Por outro lado, as CNNs exploram o fato da entrada ser uma matriz estruturada de dados para otimizar sua arquitetura. Diferente de uma ANN, os neurônios de suas camadas são organizados em largura, altura e profundidade. Vale ressaltar que o termo profundidade não se refere ao número total de camadas em uma rede. No CIFAR-10, por exemplo, o volume teria dimensões de  $32 \times 32 \times 3$  por possuir 3 canais de cores. Além disso, os neurônios em uma camada estão conectados apenas a uma pequena região da camada anterior, em vez de todos os neurônios.

Uma CNN é composta por vários tipos de camadas, que são descritas nas próximas seções: camada convolucional, camada de agrupamento (*pooling*) e camada totalmente conectada. O empilhamento dessas camadas formam uma arquitetura CNN (GOODFELLOW *et al.*, 2016). A Figura 13 mostra um exemplo da arquitetura de uma CNN com essas camadas.

Figura 13 – Arquitetura de uma CNN tradicional.



Fonte: elaborado pelo autor.

### 2.4.1 Camada convolucional

A camada convolucional é o bloco de construção central de uma CNN responsável por maior parte do trabalho computacional. Seus parâmetros consistem em um conjunto de filtros que podem ser aprendidos. Cada filtro é pequeno espacialmente (ao longo da largura e da altura), mas se estende por toda a profundidade do volume de entrada. Conforme o filtro é convoluído com esse volume de entrada, um mapa de ativação bidimensional que fornece as respostas desse filtro em cada posição espacial é produzido. Intuitivamente, a rede aprenderá filtros que se ativam com a presença de padrões visuais. Esses mapas de ativação são empilhados ao longo da dimensão de profundidade e produzem o volume de saída (GOODFELLOW *et al.*,

2016). A convolução discreta entre duas funções  $f$  e  $g$  é definida como

$$(f * g)(x) = \sum_t f(t)g(x+t). \quad (2.16)$$

Para sinais bidimensionais consideramos as convoluções 2D

$$(\mathbf{W} * \mathbf{I})(i, j) = \sum_{m,n} \mathbf{W}(m, n) \mathbf{I}(i+n, j+m), \quad (2.17)$$

em que  $\mathbf{W}$  é um *kernel* (ou filtro) de convolução aplicado a um sinal bidimensional  $\mathbf{I}$ .

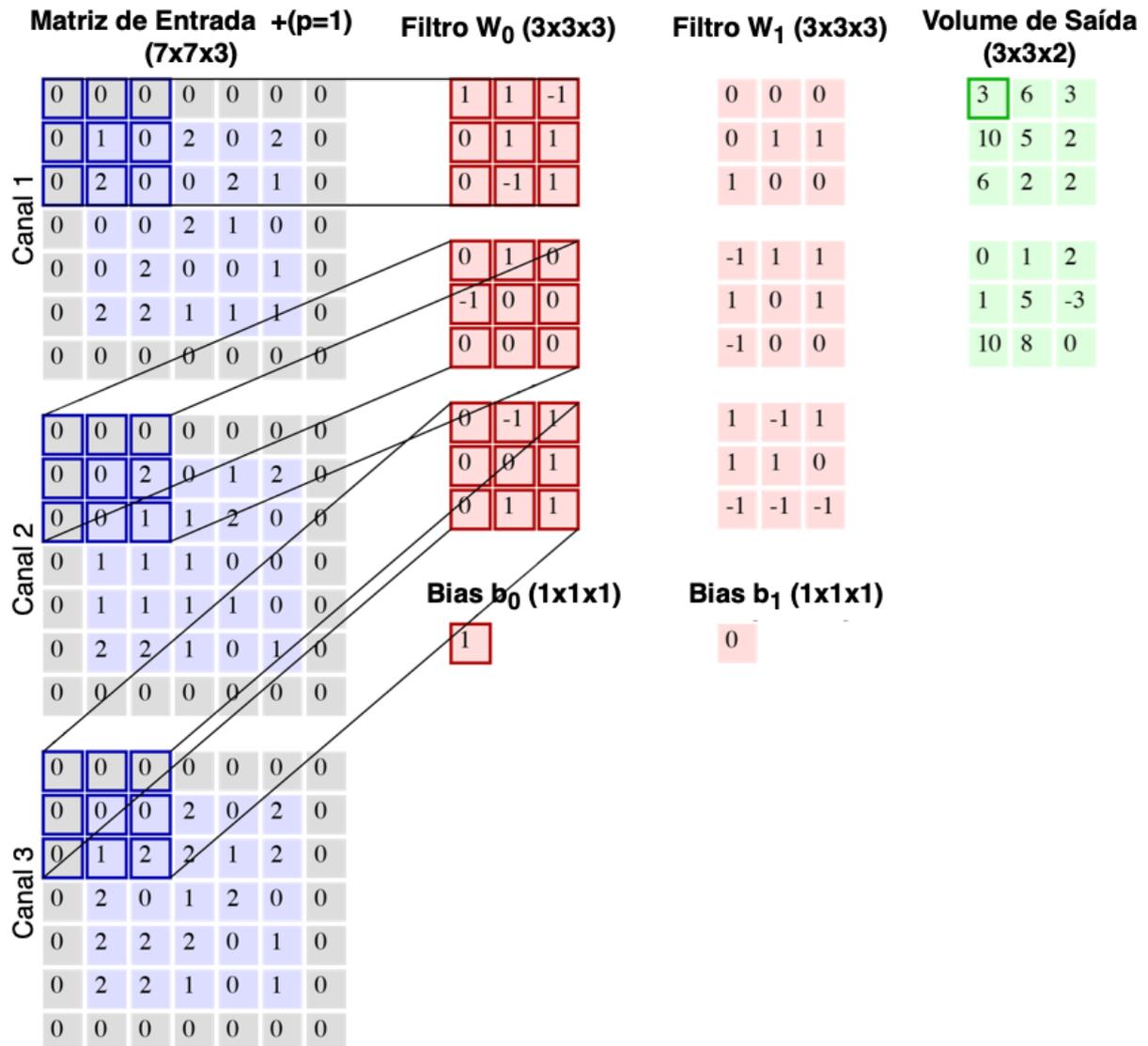
A convolução 2D consiste em mover um *kernel* sobre a matriz estruturada de dados. Em cada posição, obtém-se a convolução entre o *kernel* e a região da matriz que é calculada. Então, o *kernel* é movido por um número  $s$  de células,  $s$  é chamado de passo (*stride*). Quando o passo é pequeno, informações redundantes são obtidas. Às vezes, um preenchimento de zero (*zero padding*) também é adicionado, que é uma margem de tamanho  $p$  contendo valores zero ao redor da matriz para controlar o tamanho da saída. Suponha que  $C_{i+1}$  *kernels* são usados, cada um com tamanho  $w \times w$ . Considere ainda que o tamanho da matriz de entrada é  $L_i \times A_i \times C_i$ , onde  $L_i$  é a largura,  $A_i$  a altura e  $C_i$  o número de canais. Nesse caso, o volume de saída é  $L_{i+1} \times A_{i+1} \times C_{i+1}$ , onde  $C_{i+1}$  corresponde ao número de *kernel* considerados, e

$$\begin{aligned} L_{i+1} &= \frac{L_i - w + 2p}{s} + 1, \\ A_{i+1} &= \frac{A_i - w + 2p}{s} + 1. \end{aligned} \quad (2.18)$$

A Figura 14 exemplifica a execução da operação de convolução. Para facilitar a compreensão, a profundidade dos volumes são apresentados em fatias. A matriz de entrada possui dimensão  $(5 \times 5 \times 3)$  e está destacada em azul. São usados dois filtros na camada convolucional de tamanho  $3 \times 3$  (matriz vermelha). Além disso, os valores do passo e preenchimento são  $s = 2$  e  $p = 1$ . Logo, o tamanho volume de saída é  $(5 - 3 + 2)/2 + 1 = 3$ . O exemplo mostra que cada elemento é calculado multiplicando-se elemento a elemento da entrada destacada pelo filtro, somando-os e, em seguida, compensando o resultado pelo *bias*.

<sup>2</sup> Disponível em: <<https://cs231n.github.io/convolutional-networks/>>. Acesso em 22 de setembro de 2020.

Figura 14 – Ilustração da operação de convolução. A matriz de entrada está destacada em azul, os filtros da camada convolucional em vermelho e o volume de saída resultante da convolução em verde.



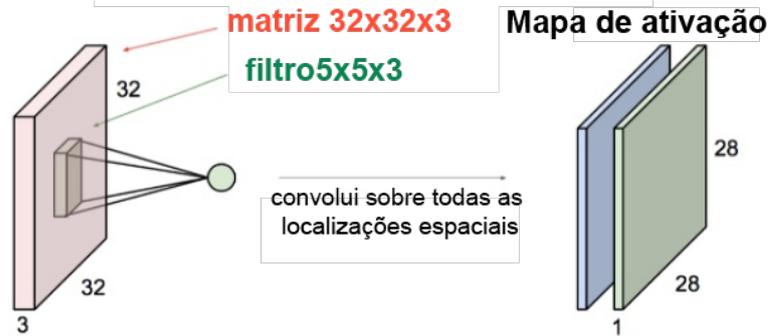
Fonte: Adaptado de CS231n Convolutional Neural Networks for Visual Recognition.<sup>2</sup>

Se a matriz tem 3 canais e se  $W_l$  ( $l = 1, \dots, C_{i+1}$ ) denota kernels  $5 \times 5 \times 3$ , a convolução da matriz  $I$  com o kernel  $W_l$  corresponde à Equação 2.19.

$$W_l * I(i, j) = \sum_{c=0}^2 \sum_{n=0}^4 \sum_{m=0}^4 W_l(n, m, c) I(i+n-2, i+m-2, c). \quad (2.19)$$

De maneira mais geral, para imagens com canais  $C_i$ , o kernel é  $(w, w, C^i, C^{i+1})$ , onde  $C^{i+1}$  é a profundidade da saída, ou seja, o número de kernels considerado. Na Figura 15, por exemplo, o kernel é  $(5, 5, 3, 2)$ . O número de parâmetros associado a um kernel de forma  $(w, w, C^i, C^{i+1})$  é  $(w \times w \times C^i + 1) \times C^{i+1}$ .

Figura 15 – Mapa de ativação gerado por uma camada convolucional com 2 filtros.



Fonte: elaborado pelo autor.

As operações de convolução são combinadas com uma função de ativação  $\phi(\cdot)$ , geralmente a função de ativação *Relu*. Para um *kernel*  $\mathbf{W}$  de tamanho  $w \times w$  e um região  $\mathbf{x}$  do mesmo tamanho da imagem, a ativação é obtida deslizando a janela  $w \times w$  e computando  $z(\mathbf{x}) = \phi(\mathbf{W} * \mathbf{x} + b)$ , onde  $b$  é o *bias*.

É na camada convolucional que ocorre o aprendizado de fato por meio da adequação dos filtros para cada tarefa específica. Outra vantagem é que várias camadas convolucionais podem ser consideradas, assim, a saída de uma convolução torna-se a entrada da próxima.

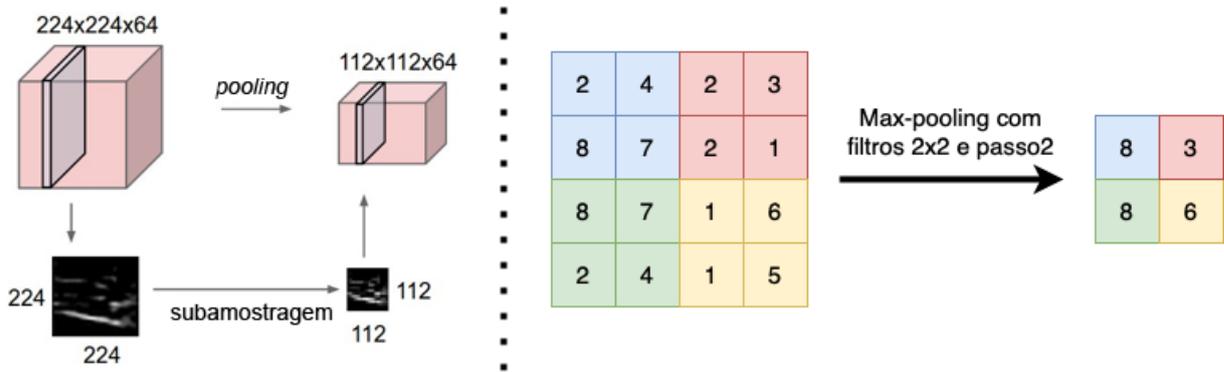
#### 2.4.2 Camada de agrupamento

Outra camada importante nas CNNs são as de agrupamento (*pooling*) também conhecida como subamostragem. Essa camada permite reduzir a dimensão, ao calcular a média ou o máximo em regiões da imagem.

Como as camadas convolucionais, a camada de *pooling* age em pequenos trechos da imagem e podem ter passos (*strides*). Se for calculado o valor máximo de regiões  $2 \times 2$  e considerando um passo  $s = 2$ , a largura e altura da imagem é dividida por 2 como mostra a Figura 16. Vale ressaltar que também é possível reduzir a dimensão da camada convolucional com um passo maior do que 1 e sem preenchimento de zero. Outra vantagem da camada de *pooling* é tornar a rede menos sensível a pequenas translações das imagens de entrada.

De forma mais geral, a camada de *pooling* aceita um volume de tamanho  $L_i \times A_i \times C_i$ , requer dois hiperparâmetros (sua extensão espacial e o passo) e produz um volume de tamanho

Figura 16 – Ilustração do efeito da camada de *pooling* com operador máximo.



Fonte: Adaptado de *CS231n Convolutional Neural Networks for Visual Recognition*.<sup>2</sup>

$L_{i+1} \times A_{i+1} \times C_{i+1}$ , onde

$$L_{i+1} = \frac{L_i - w}{s} + 1,$$

$$A_{i+1} = \frac{A_i - w}{s} + 1, \quad (2.20)$$

$$C_{i+1} = C_i.$$

### 2.4.3 Camada totalmente conectada

Os neurônios em uma camada totalmente conectada têm conexões para todas as ativações na camada anterior, como visto nas ANNs. Suas ativações podem, portanto, ser calculadas com uma multiplicação de matriz juntamente com o *bias*.

A única diferença entre as camadas totalmente conectadas e as convolucionais é que os neurônios dessa estão conectados apenas a uma região local na entrada, e que muitos deles compartilham parâmetros. No entanto, os neurônios em ambas as camadas ainda computam produtos escalares, então sua forma funcional é idêntica. Portanto, uma conversão entre essas camadas é possível (GOODFELLOW *et al.*, 2016).

## 2.5 Considerações finais

Neste capítulo, a função e a estrutura do coração, bem como os aspectos do eletrocardiograma são descritos. Este capítulo também apresenta uma série de doenças cardiovasculares que podem ser diagnosticadas com o ECG. Além disso, na seção sobre o aprendizado profundo é elucidado o funcionamento das redes neurais convolucionais.

### 3 METODOLOGIA

Este capítulo apresenta a metodologia adotada na pesquisa. Inicialmente, investigou-se a melhor combinação de extratores de características e classificadores utilizando a abordagem clássica para classificação de arritmias cardíacas em sinais de eletrocardiograma, Seção 3.1. Os desafios encontrados neste primeiro grupo de estudo motivou a criação de um novo método que através de uma abordagem de regressão com Aprendizado Profundo é capaz de simultaneamente classificar arritmias cardíacas, detectar e segmentar os batimentos em sinais de eletrocardiograma. A formulação matemática do método responsável pelo segundo grupo de experimentos é explicada na Seção 3.2. As especificações aderidas para caracterização das arritmias cardíacas com objetivo de atender os padrões clínicos são elucidadas na Seção 3.3. Em seguida, apresenta-se a base de dados adotada na avaliação do método proposto durante a detecção e segmentação dos batimentos cardíacos no ECG, Seção 3.4. Por fim, a Seção 3.5 explica as métricas de avaliação utilizadas na validação dos métodos.

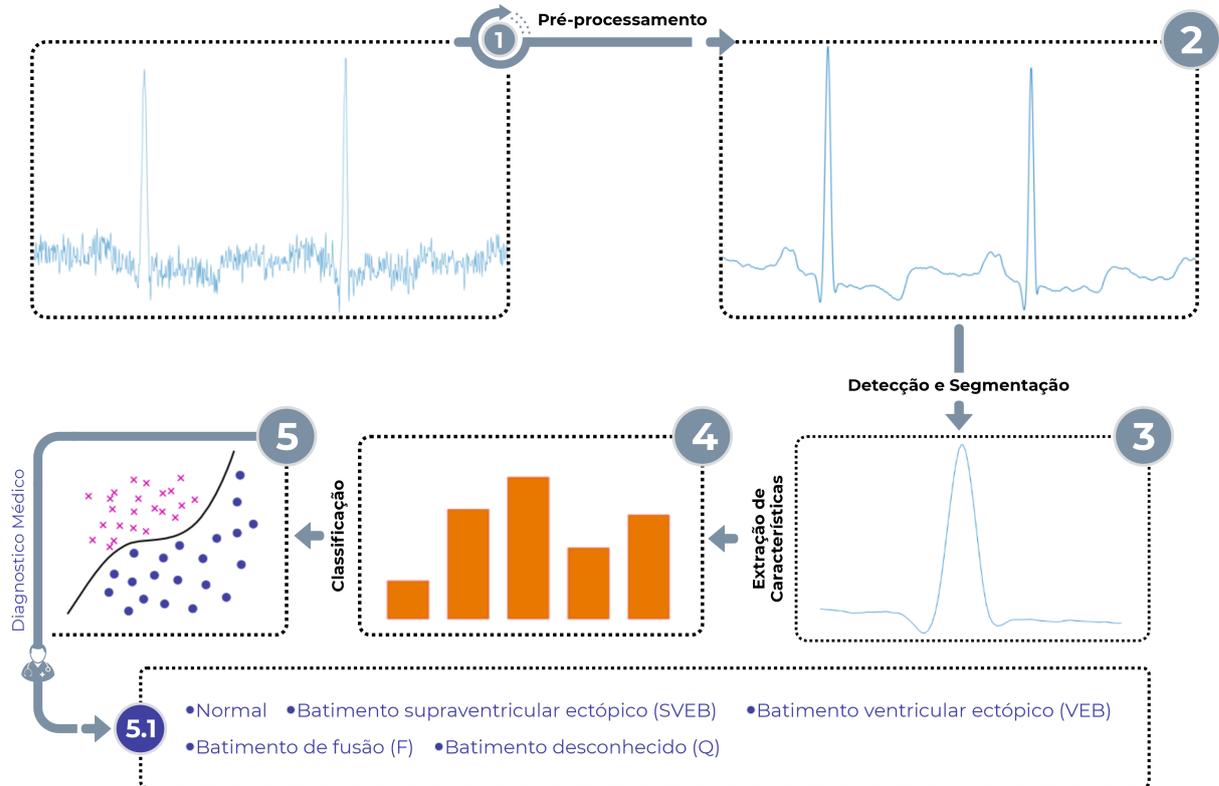
#### 3.1 Abordagem clássica para classificação de arritmias cardíacas usando sinais de ECG

Os sistemas tradicionais de classificação de arritmias cardíacas em sinais de ECG utilizam basicamente as seguintes etapas: aquisição, pré-processamento, detecção, segmentação, extração de características e classificação. A Figura 17 mostra o fluxograma da abordagem clássica adotada nestes sistemas. Após a aquisição dos dados, a etapa de pré-processamento filtra as frequências de ruído e de artefatos que possam interferir na banda de frequência do ECG. Em seguida, é preciso detectar e segmentar cada batimento do sinal de ECG. Mesmo com os batimentos em destaque, muitas informações clinicamente relevantes não estão evidente no sinal de ECG. Desta forma, extratores de características precisam ser utilizados para obter os atributos que serão usados como entrada dos classificadores.

As etapas de extração de características e classificação são as mais desafiadoras e podem afetar significativamente o resultado final. Diante disso, a pesquisa inicia em busca da resposta para seguinte pergunta: **Qual a combinação de extrator de características e classificador é mais adequada para classificação de arritmia cardíaca em sinais de ECG?**

Os estudos que adotam as etapas citadas anteriormente são organizados em dois tipos principais de paradigmas de avaliação, inter-paciente e intra-paciente, como são descritos na Seção 3.3.2. As comparações entre os resultados das pesquisas que não utilizam o paradigma

Figura 17 – Fluxograma das etapas adotadas na abordagem clássica para classificar arritmias cardíacas em sinais de ECG.



Fonte: elaborado pelo autor.

inter-paciente são mais complexas, pois não há uma padronização entre os protocolos de avaliação. Ademais, estes estudos podem não apresentar uma relevância clínica pela divergência do contexto real.

Diante disso, Luz *et al.* (2013) reavaliaram alguns extratores de características com taxa de acerto próxima de 100%, mas que não seguiam nenhum protocolo de seleção de batimentos cardíacos. Estes métodos baseiam-se na Transformada Wavelet Discreta (*Discrete Wavelet Transform, DWT*), Análise de Componentes Independentes (*Independent Component Analysis, ICA*), Análise de Componentes Principais (*Principal Component Analysis, PCA*) e informações sobre o intervalo RR. Os extratores considerados e comparados com os métodos da abordagem clássica desta tese são:

- a) morfologia do sinal e intervalos RR (CHAZAL *et al.*, 2004);
- b) DWT (GÜLER; ÜBEYLI, 2005);
- c) DWT (SONG *et al.*, 2005);
- d) DWT, intervalo RR e energia do sinal (YU; CHEN, 2007);
- e) DWT, ICA e intervalo RR de (YU; CHOU, 2008);
- f) DWT, ICA, PCA e intervalo RR (YE *et al.*, 2010).

Assim, a primeira etapa de estudos nesta tese investiga diferentes categorias de extratores de características e classificadores e os compara com os principais trabalhos existentes para tarefa de interesse.

### 3.1.1 Métodos de extração de características

Dentre as técnicas de extração de características existentes na literatura, adotou-se diferentes categorias de métodos consolidados com intuito de escolher a configuração mais adequada para classificação de arritmias cardíacas em sinais de ECG. Avaliou-se a transformada de Fourier, Goertzel, Estatísticas de Alta Ordem (*Higher-Order Statistics*, HOS), Matriz de Co-ocorrência Estrutural (*Structural Co-occurrence Matrix*, SCM) e SCM-Fourier. O pequeno número de atributos gerados pelos três últimos métodos pode melhorar a etapa de classificação. Além disso, a SCM foi usada pela primeira vez com sinais de ECG em Marinho *et al.* (2019).

A transformada de Fourier é uma ferramenta projetada para descrever um sinal no domínio da frequência e destacar as características importante nas variações de frequência do sinal. Existem várias aplicações desta transformada, como filtragem, reconstrução e compressão de sinais, bem como extração de características para reconhecimento de padrões (ERSOY, 1994). A transformada de Fourier é amplamente utilizada em aplicações práticas e acadêmicas, viável para uso computacional através do algoritmo Transformada Rápida de Fourier (*Fast Fourier Transform*, FFT) (BRIGHAM; BRIGHAM, 1988).

O algoritmo Goertzel foi proposto como uma forma de calcular os coeficientes da transformada discreta de Fourier. É um algoritmo recursivo que realiza operações aritméticas com um único valor em cada interação, facilitando o cálculo das harmônicas. A vantagem de usar este algoritmo é que ele é numericamente mais eficiente para calcular valores individuais de harmônicas, o que o torna aplicável a aplicações em tempo real (GOERTZEL, 1958).

O uso da Curtose foi proposto pela primeira vez por Dwyer (1983) como uma ferramenta estatística para indicar componentes não Gaussianos em um sinal. Em seguida, a teoria foi reformulada por Capdevielle *et al.* (1996) e descrita a partir do momento estatístico normalizado de quarta ordem para criação das Estatísticas de Alta Ordem (*Higher-Order Statistics*, HOS). A literatura mostra o uso dos momentos estatísticos normalizados de terceira ordem (Assimetria) combinados com a Curtose para auxiliar na análise vibracional de rolamentos (MARTIN; HONARVAR, 1995; SAMANTA; AL-BALUSHI, 2003) em conjunto com técnicas de aprendizado de máquina. A vantagem do uso deste método em sinais está na sua insensibilidade ao ruído do

tipo Gaussiano ao usar momentos de ordem maiores que dois (MENDEL, 1991). Esta afirmação foi posteriormente confirmada por Li *et al.* (2008).

A Matriz de Co-ocorrência Estrutural (*Structural Co-occurrence Matrix*, SCM) é um método de análise estrutural de sinais discretos que se baseia em estatísticas de co-ocorrência. A premissa deste método é que existem conexões entre estruturas de baixo nível de dois sinais discretos  $N$ -dimensional. Sua principal característica é apresentar um conhecimento prévio dos sinais em análise, com ênfase na detecção de detalhes. Sua saída é um histograma bidimensional, onde a SCM fornece as co-ocorrências entre as estruturas dos sinais de entrada (RAMALHO *et al.*, 2016). Uma representação do sinal de entrada será comparada a outro sinal, que corresponde ao sinal de entrada após a aplicação de um filtro pré-selecionado. Após esta comparação, gera-se uma matriz com as diferenças estruturais entre os sinais. Em seguida, os atributos são extraídos do histograma bidimensional. Além disso, Ramalho *et al.* (2016) propôs seis atributos a serem extraídos de qualquer sinal.

Na literatura, o uso de SCM foi relatado em tarefas de classificação de imagens (SILVA *et al.*, 2017; PEIXOTO; FILHO, 2018). A primeira aplicação da SCM como extrator de características em sinais de ECG foi em Marinho *et al.* (2019). Filho *et al.* (2018) propuseram uma variação da SCM para detecção automática de lesões cutâneas em imagens. Seu funcionamento é semelhante a SCM mencionada anteriormente; entretanto, um dos sinais de entrada é o sinal bruto reconstruído usando as harmônicas de Fourier que têm 1% da componente DC. Nascimento *et al.* (2020) foram os primeiros a usar esta variação da SCM em sinais de ECG.

### **3.1.2 Parâmetros para segmentar os batimentos na abordagem clássica**

A partir das anotações do banco de dados MIT-BIH, o sinal foi segmentando usando o mesmo intervalo usado em (YU; CHEN, 2007) para extrair as características posteriormente. Yu e Chen (2007) usou uma janela de 64 amostras, onde o pico R é o centro. Como a taxa de amostragem é de 360Hz e o tempo entre as amostras é de 2,8ms, o intervalo adotado para segmentar o sinal foi  $[\text{batimento} - 32, \text{batimento} + 31]$ . É importante salientar que na literatura, cada uma das técnicas clássicas de extração de características usa uma faixa diferente para segmentar o sinal.

### 3.1.3 Métodos de classificação

De forma similar aos extratores de características, adotou-se diferentes tipos de classificadores com objetivo de encontrar a melhor configuração para a tarefa classificação de arritmias cardíacas em sinais de ECG. Dentre as categorias de técnicas existentes, utilizou-se o classificador Bayesiano, o qual se baseia em probabilidade e estatística, o que usa a Teoria de Aprendizado Estatístico, Máquina de Vetores de Suporte (*Support Vector Machine*, SVM), a rede neural artificial Perceptron Multicamada (*Multilayer Perceptron*, MLP) e o classificador que adota a teoria de Grafos, Floresta de Caminhos Ótimos (*Optimum-Path Forest*, OPF).

A Perceptron Multicamada (*Multilayer Perceptron*, MLP) é uma ferramenta poderosa na modelagem de mapeamentos de entrada e saída normalmente encontradas em problemas de aproximação de função e classificação de padrões. A MLP consiste em um sistema de neurônios simples interconectados, representando um mapeamento não linear entre um vetor de entrada e um vetor de saída. Os pesos ligam os neurônios (sinapses) e os sinais de saída, que são a soma das entradas ponderadas e modificados por uma função de ativação (HAYKIN, 2001).

Os classificadores Bayesianos usam cálculos probabilísticos entre as classes presentes nos dados para classificar novas amostras como pertencentes à classe mais provável. Levando em conta as diferentes quantidades de amostras por classe e os diferentes valores de variância das mesmas, a teoria da decisão estatística proposta por Fisher (1922) maximiza a probabilidade *a posteriori* e classifica uma amostra desconhecida como pertencente à classe mais provável. No classificador Bayesiano, assume-se que os dados se comportam como uma função de densidade de probabilidade Gaussiana. Quando utiliza-se uma matriz de covariância diagonal, considera-se que todos os atributos são estatisticamente independentes. Esta abordagem é adotada no classificador Naive-Bayes (RUSSELL; NORVIG, 2016).

O classificador Floresta de Caminhos Ótimos (*Optimum-Path Forest*, OPF) modela o problema de reconhecimento de padrões como partições de um grafo em um determinado espaço de características. Neste modelo, os vértices representam atributos dos vetores e as arestas conectam todos os pares de vértices, formando um grafo completo (PAPA *et al.*, 2009). As arestas entre as amostras contêm pesos que variam com as distâncias estabelecidas entre seus respectivos vetores atributos. Luz *et al.* (2013) usaram o OPF para classificação de arritmias cardíacas.

A partir do princípio de minimização de risco estrutural, o processo de otimização do Máquina de Vetores de Suporte (*Support Vector Machine*, SVM) busca estabelecer uma função

de separação ponderando entre generalização e o sobreajuste (*overfitting*) durante a classificação (VAPNIK, 1999; CORTES; VAPNIK, 1995). A fim de resolver problemas de não linearidade na superfície de decisão, o produto interno entre os pesos e vetores podem ser substituídos por uma função de *kernel*. Dentre as várias funções de *kernel*, a Função de Base Radial (*Radial Basis Function*, RBF) é uma das mais utilizadas (WANG *et al.*, 2004).

Após a etapa de extração de características, oito conjuntos de dados foram criados a partir dos quatro extratores considerando a divisão com 5 e 3 classes. Todos os bancos de dados foram igualmente pré-processados. O conjunto de dados de treinamento foi normalizado (média de zero e variância unitária) e os conjuntos de dados de teste também foram normalizados da mesma forma que o conjunto de treinamento. As classes foram divididas de forma equilibrada em conjuntos de dados de treinamento e teste de acordo com as recomendações da norma ANSI/AAMI EC57 descrita na Seção 3.3.2. Nessa tese, para cada uma dos conjuntos de dados provenientes da etapa de extração de características, particionou-se aleatoriamente os padrões em dois grupos, onde oitenta por cento dos dados foram utilizados na etapa de treinamento e o restante para classificação. Esse procedimento foi repetido dez vezes.

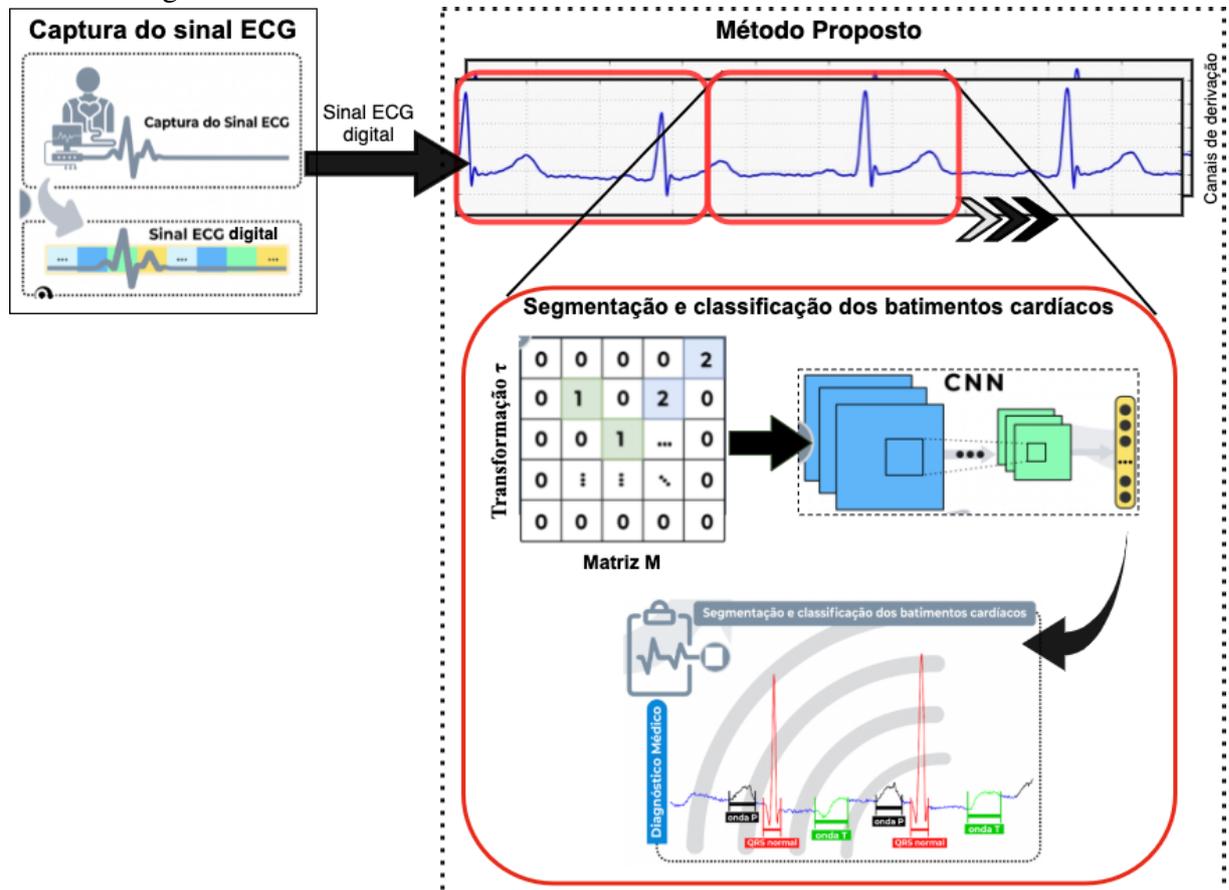
Os hiperparâmetros para os classificadores SVM e MLP foram escolhidos através das técnicas de validação cruzada com 10 *folds* com busca aleatória (BERGSTRA; BENGIO, 2012). Utilizou-se o *kernel* RBF no SVM com intervalo  $[2^{-15}, 2^{33}]$  para  $\gamma$  e  $[2^{-5}, 2^{15}]$  para o  $C$ . A MLP foi treinada com o método de otimização Adam (KINGMA; BA, 2015) e o intervalo de neurônios usados na camada oculta foi de 2 a 500. A função densidade de probabilidade Gaussiana foi aplicada no classificador Naive-Bayes.

### 3.2 Formulação matemática do método proposto

Diante dos desafios encontrados com os métodos da abordagem clássica, um novo método para classificar arritmias cardíacas e, ao mesmo tempo, detectar e segmentar os batimentos em sinais de eletrocardiograma (ECG) por meio de regressão com Aprendizado Profundo foi proposto. Diferente das técnicas de processamento digital de sinais existentes, o método apresentado nesta tese usa um único modelo para simultaneamente realizar as tarefas supracitadas. Outro diferencial é a habilidade de classificar a anomalia cardíaca sem a etapa explícita de extração de características presente nas abordagens tradicionais. O método mostrou boa capacidade de generalização ao ser avaliado em bases distintas com condições fisiológicas e frequências de amostragem diversificadas.

A Figura 18 apresenta as etapas empregadas pelo método para obtenção do diagnóstico médico a partir do sinal ECG digitalizado. O método está dividido em três etapas principais: pré-processamento, regressão para computação das janelas e pós-processamento. Durante o pré-processamento, o sinal de entrada é dividido em trechos e realizada a transformação  $\tau$  em cada um deles. Na etapa seguinte, cada uma das matrizes  $\mathbf{M}$  obtidas anteriormente são utilizadas como entrada de uma CNN cuja saída é um tensor com valores de uma regressão. Esses valores representam as janelas segmentadas e o tipo de batimento. Por fim, a etapa de pós-processamento faz alguns ajustes finos para filtrar saídas duplicadas.

Figura 18 – Etapas adotadas no método para obtenção do diagnóstico médico usando sinal do eletrocardiograma.



Fonte: elaborado pelo autor.

O método reformula as tarefas de segmentação e classificação das ondas como um único problema de regressão a partir das informações no domínio no tempo. Observa-se também a capacidade de aprender representações generalistas dos batimentos cardíacos. Por conta dessa característica, ele se torna menos propenso a falhar quando aplicado a novos domínios ou com

entradas incomuns.

### 3.2.1 Pré-processamento do sinal de entrada

A etapa de pré-processamento tem como objetivo preparar o sinal de entrada para o método de Aprendizado Profundo que aprenderá os batimentos. Dado que serão aprendidas características morfológicas do sinal e buscando meios de paralelizar o processamento, adotou-se uma estratégia para dividir o sinal em trechos. Além disso, uma transformação em cada trecho é feito para adequar as entradas para CNN.

Verificou-se por meio de vários experimentos que o número de pontos (atividade elétrica do coração digitalizada) dentro de cada fragmento pode ser entre a taxa de amostragem e a seguinte relação,  $\text{taxa de amostragem}/0,45$ . Ou seja, para a base de dados MIT-BIH cuja taxa de amostragem é 360Hz, teremos trechos de 360 pontos até 800 pontos. Enquanto que para a base de dados QTDB com até 250Hz, teremos 250 até 556 pontos em cada trecho. A Figura 19 apresenta trechos do sinal no registro 100 para as bases de dados QT e MIT-BIH.

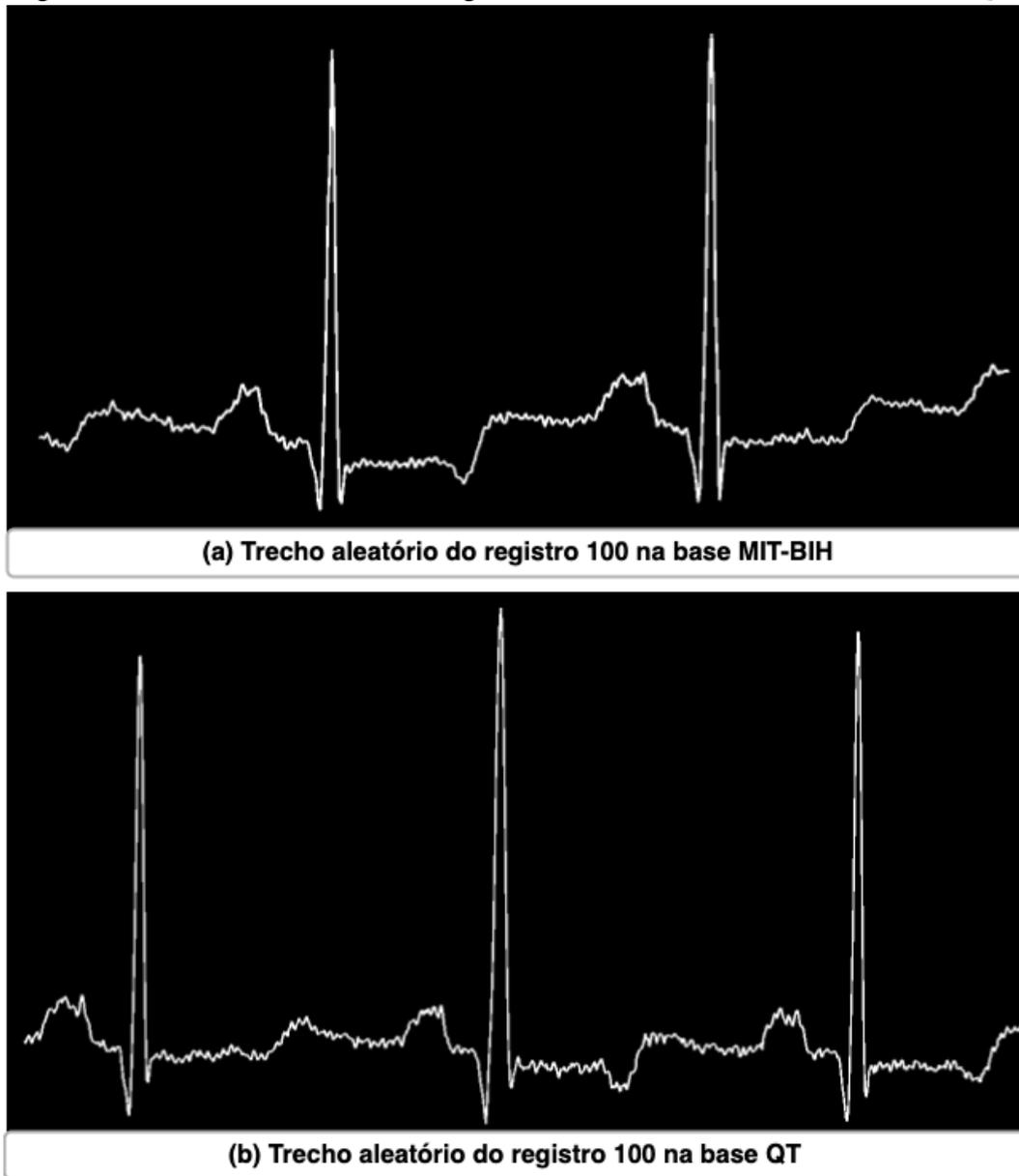
A exploração de relações complexas entre as ondas do sinal ECG pode não ser suficiente diretamente no espaço original dos dados de entrada. Em alguns casos, é importante empregar um mapeamento para outro espaço que represente de modo mais natural a estrutura dos batimentos cardíacos. Isso permite uma melhor exploração das características morfológicas do sinal.

Na Máquina de Vetores de Suporte (*Support Vector Machine*, SVM), por exemplo, o Truque do *Kernel* possibilita que o espaço original seja mapeado em um espaço de produto escalar de alta-dimensão chamado espaço de características, onde os dados podem ser linearmente separáveis (MURTY; RAGHAVA, 2016). Uma abordagem para representar o sinal ECG em duas dimensões adotando mais de um canal é chamada de Vetorcardiograma (VCG) (MOODY; MARK, 2001).

O método proposto nesta tese transforma as amostras das atividades elétricas do coração espaçadas no tempo,  $s(t)$ , em um tensor que permite, por exemplo, a combinação de mais de um canal de derivação. No novo espaço criado, as informações morfológicas dos batimentos são realçadas após o uso da rede neural convolucional. Além disso, vale ressaltar a habilidade das CNNs em classificar imagens e a semelhança do espaço criado no método proposto com essas matrizes estruturadas de dados.

A Equação 3.1 apresenta a matriz  $\mathbf{M} \in \mathbb{R}^{N \times T \times C}$  criada a partir do mapeamento

Figura 19 – Trechos aleatórios do registro 100 nas base de dados MIT-BIH e QT.



Fonte: elaborado pelo autor.

$\tau : \mathbb{R} \rightarrow \mathbb{R}^C$  sobre o sinal  $s(t)$ , tal que  $s : \mathbb{R} \rightarrow \mathbb{N}$  e  $1 \leq t \leq T$ , cujo  $T$  representa a última medição do sinal.  $C$  indica a quantidade de canais de derivação. No exemplo abaixo, com  $C = 1$ , temos o valor da atividade elétrica do coração no momento  $t$ , denotado por  $s(t)$ . Dessa forma, os menores

valores do sinal ficarão na primeira linha,  $M_{1,*}$ , e os maiores na última linha,  $M_{*,N}$ :

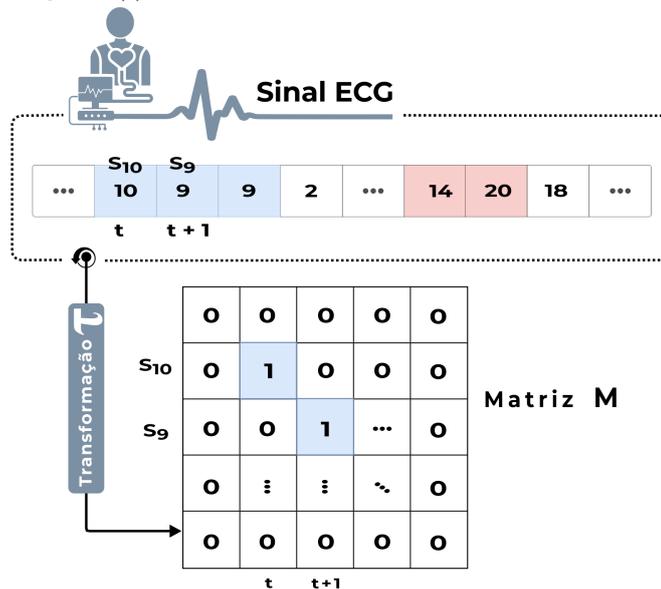
$$M_{i,j,1} = \begin{cases} 1, & \text{se } s(t) \neq 0; \\ 0, & \text{caso contrário,} \end{cases} \quad (3.1)$$

tal que  $i = t$  e

$$j = \delta(t) = \left\lfloor \frac{s(t) - \min(s)}{\max(s) - \min(s)} \times (T - 1) \right\rfloor + 1.$$

A Figura 20 exemplifica a criação da matriz  $\mathbf{M}$ . Os valores marcados pela cor azul no vetor representam um batimento cardíaco da classe 1, enquanto os valores na cor vermelha são da classe 2. Após a transformação  $\tau(\cdot)$ , a matriz  $\mathbf{M}$  receberá o valor 1 na linha 10 e coluna  $t$  por conta do valor de  $s_{10}$  que está na posição temporal  $t$  do sinal ECG. Esse raciocínio é efetuado sobre os demais pontos do sinal.

Figura 20 – Exemplo da criação da matriz  $\mathbf{M}$  após a transformação  $\tau(\cdot)$ .



Fonte: elaborado pelo autor.

### 3.2.2 Processamento das janelas usando regressão com Aprendizado Profundo

Cada matriz  $\mathbf{M}$  gerada na etapa de pré-processamento é submetida a uma CNN cuja saída é um tensor com valores de regressão representando as localizações da onda dentro do sinal avaliado. Além disso, cada onda também é rotulada usando a mesma rede, conseguindo dessa forma classificar os tipos de arritmias cardíacas.

A motivação para o uso de uma CNN para aprender as ondas e ao mesmo classificar os tipos de arritmias cardíacas surgiu a partir da estratégia adotada em técnicas de detecção de objetos em imagens que permite identificá-los e classificá-los com o mesmo modelo. Assim, o método proposto nesta tese consiste em realizar a segmentação e classificação dos batimentos no sinal ECG usando uma abordagem como da *You Only Look Once*, YOLO (REDMON *et al.*, 2016).

Embora o método de Aprendizado Profundo adotado baseou-se na YOLO, outras técnicas de CNN que tenham como entrada uma matriz como a descrita na transformação  $\tau(\cdot)$  e uma saída representada por uma regressão onde os valores representam a disposição das ondas e há uma classe para identificar o tipo de batimento também pode ser utilizados. A seguir, serão explicados a matemática empregada para o aprendizado das ondas juntamente com a classificação dos tipos de batimentos.

Após a transformação  $\tau(\cdot)$ , o método particiona a matriz  $\mathbf{M}$  em  $K$  regiões contíguas e disjuntas, isto é,  $\mathcal{K}_i \cap \mathcal{K}_j = \emptyset$ , para  $i \neq j$ , com  $i, j = 1, 2, \dots, K$ . Se o centro de um das ondas do ECG estiver em um das regiões, a mesma é responsável por caracterizar a onda. Essa divisão é inspirada no papel usado para escrita do electrocardiograma e da abordagem empregada pelo método de detecção de objetos em imagens YOLO (REDMON *et al.*, 2016).

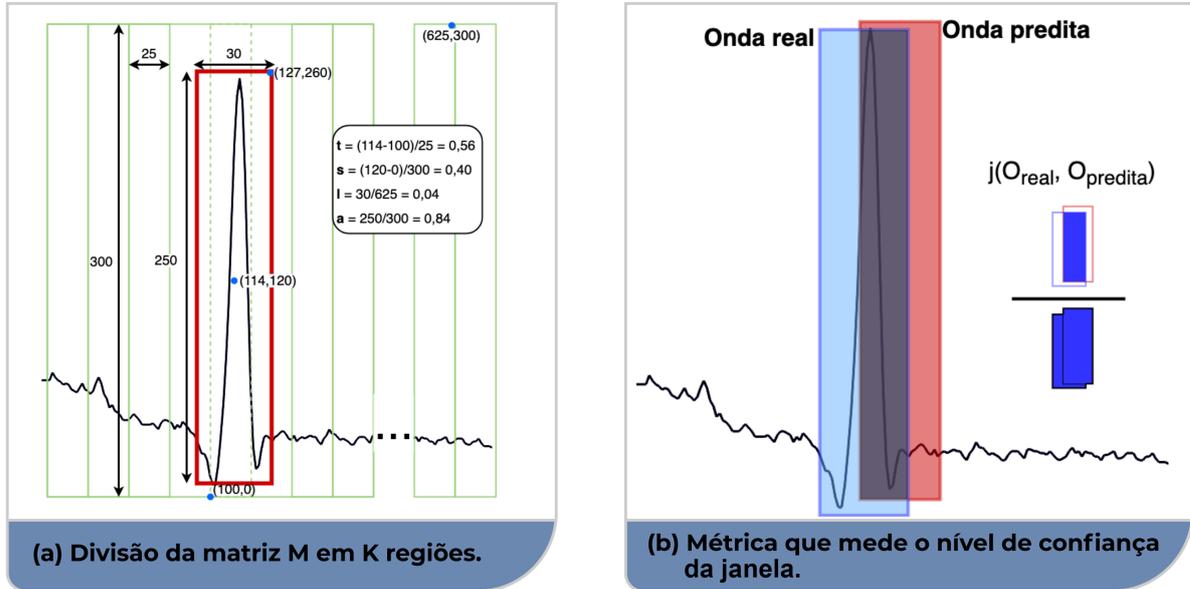
A Figura 21(a) exemplifica as partições criadas pelo método. A coordenada (114, 120) representa o centro da onda QRS, enquanto a largura e a altura do batimento são indicadas por  $l = 30$  e  $a = 250$ . Esses valores são normalizados e apresentados no quadro ao lado. A janela da onda está destacada pela linha vermelha, e as regiões pelas linhas verdes. A região com a linha tracejada é a responsável por caracterizar a onda.

Cada região prediz  $J$  janelas e pontuações que indicam o nível de confiança para essas janelas. Essas pontuações refletem a confiança do modelo em que a janela contém uma onda no ECG e também quão preciso é a classificação para a janela. Formalmente, defini-se o nível de confiança como  $P(O) * j(O_{real}, O_{predita})$ . Se nenhuma onda existir na região, a pontuação de confiança deve ser zero,  $P(O) = 0$ . Caso contrário, espera-se que esse número seja igual à  $j(O_{real}, O_{predita})$ . A Equação 3.2 mostra como calcular o  $j(O_{real}, O_{predita})$  para uma onda do tipo  $B$ . Um exemplo dessa métrica é ilustrado na Figura 21(b).

$$j(O_{real}, O_{predita}) = \frac{|Onda_{real} \cap Onda_{predita}|}{|Onda_{real} \cup Onda_{predita}|}. \quad (3.2)$$

Cada janela é composta por cinco predições:  $s, t, l, a$ , e a porcentagem de confiança.

Figura 21 – (a) exemplo da  $K$  regiões criadas após a transformação  $\tau(\cdot)$  e (b) métrica empregada para calcular a precisão de segmentação dos batimentos cardíacos,  $j(O_{real}, O_{predita})$ .



Fonte: elaborado pelo autor.

A coordenada  $(t, s)$  representa o centro da janela da onda em relação aos limites da região. A largura do batimento é indicado por  $l$ , enquanto a diferença da localização entre o maior e menor valor do batimento na matriz  $\mathbf{M}$  é simbolizado por  $a$ . Por fim, a porcentagem de confiança é representada por  $j(O_{real}, O_{predita})$ .

Cada região também prevê  $B$  probabilidades condicionais dos tipos de onda,  $P(B_i|O)$ . Essas probabilidades são condicionadas na região que contém uma onda. A CNN prevê apenas um conjunto de probabilidades das classes por região, independentemente do número de janelas.

Durante a etapa de classificação, multiplica-se as probabilidades condicionais de cada classe e as porcentagens de confiança da janela de cada onda,

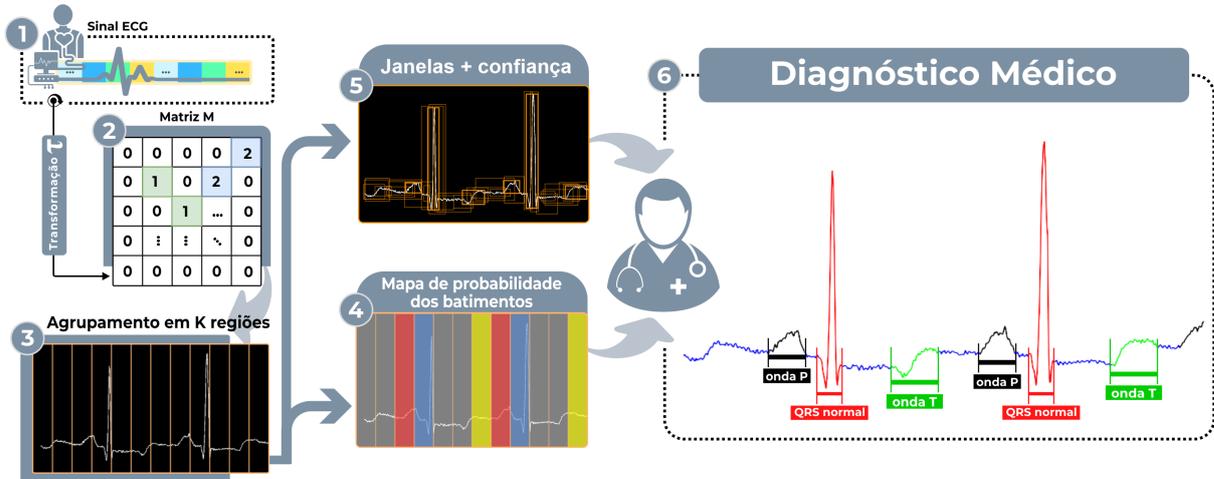
$$P(B_i|O) * (P(O) * j(O_{real}, O_{predita})) \quad (3.3)$$

$$\propto P(B_i) * j(O_{real}, O_{predita}),$$

o que nos dá pontuações de confiança específicas para cada classe. Essas pontuações codificam a probabilidade de que a classe apareça na janela e quão bem foi a segmentação da onda.

A Figura 22 mostra um esquema da arquitetura unificada para segmentação e classificação de batimentos cardíacos. A primeira camada do método transforma o sinal ECG na matriz  $\mathbf{M}$ . Posteriormente, essa matriz é dividida em  $K$  regiões. Para cada uma dessas regiões são previstas  $J$  janelas, o grau de confiança para as janelas das ondas e as probabilidades da classe  $C$ . Essas previsões são codificadas como um tensor  $K \times (J * 5 + B)$ .

Figura 22 – Diagrama da arquitetura unificada para segmentação e classificação do batimentos cardíacos.



Fonte: elaborado pelo autor.

### 3.2.2.1 Projeto e treinamento da Rede Neural Convolutacional

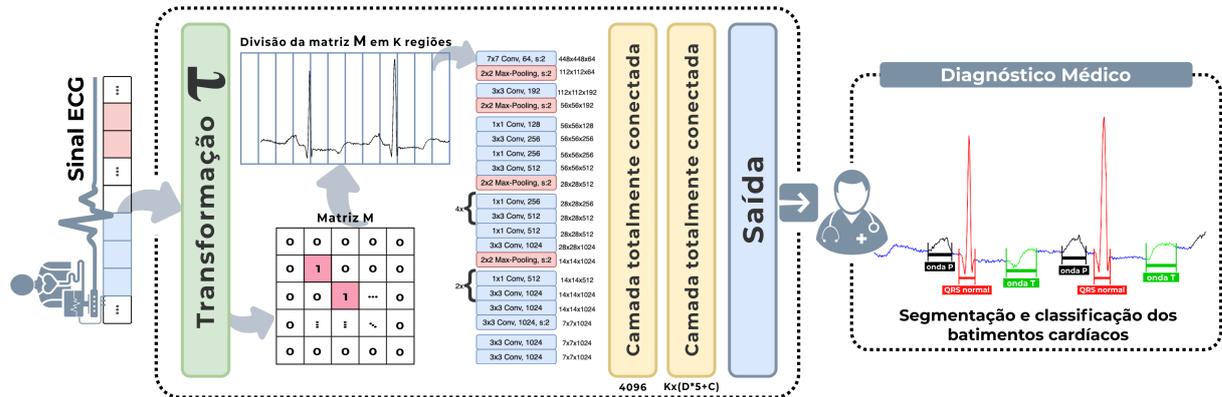
As abordagens tradicionais para reconhecimento de anomalias cardíacas em sinais de eletrocardiograma consiste em segmentar, extrair características e, finalmente, classificar os padrões. Por outro lado, as Redes Neural Convolutacional (*Convolutional Neural Network*, CNNs) detectam automaticamente os atributos relevantes sem a necessidade de extratores ou da supervisão humana.

Algumas arquiteturas de CNNs foram propostas nas últimas décadas e são amplamente utilizadas em diversos segmentos. Dentre elas, destaca-se a GoogLeNet (*Inception*) por ser a primeira a chegar perto do desempenho humano (SZEGEDY *et al.*, 2015). Diante disso, adaptou-se sua arquitetura para a tarefa simultânea de classificação de doenças cardíacas e segmentação de batimentos utilizando sinais ECG.

A rede completa é mostrada na Figura 23. Enquanto as primeiras camadas convolucionais do método extraem características do sinal, as camadas totalmente conectadas predizem as probabilidades e coordenadas dos batimentos cardíacos. Ou seja, as tarefas de segmentação e classificação simultânea das ondas foram reformuladas como único problema de regressão a partir das informações no domínio no tempo.

A camada final do modelo prediz as probabilidades de cada tipo de batimento juntamente com as janelas que representam as referências temporais no sinal ECG. A largura e as coordenadas das janelas foram normalizadas pela largura do sinal, de forma que fiquem entre 0 e 1. Além disso, adotou-se uma função de ativação linear para a camada final e todas as outras

Figura 23 – Arquitetura da rede neural convolucional adotada para segmentar e classificar simultaneamente os batimentos cardíacos do sinal ECG.



Fonte: elaborado pelo autor.

camadas usam a função *Leaky ReLU*,  $\max(0, 1x; x)$ .

A saída do modelo é otimizada pela Soma dos Quadrados dos Erros,  $SQE = \sum_i^n (e_i)^2$ . Mais especificamente, busca-se minimizar a diferença relacionada à posição da janela da onda para todas as predições ( $j = 0, \dots, J$ ) em cada região ( $i = 0, \dots, K$ ),

$$f(\cdot) = \sum_{i=0}^K \sum_{j=0}^J (t_i - \hat{t}_i)^2 + (s_i - \hat{s}_i)^2. \quad (3.4)$$

Embora a *SQE* seja simples de otimizar, ela não se alinha perfeitamente ao objetivo de maximizar a precisão média. Essa métrica não pondera entre o erro de localização e de classificação. Ademais, em alguns sinais, a região pode não conter onda. Isso leva as pontuações de confiança dessas regiões para zero, geralmente sobrecarregando o gradiente das regiões que contêm ondas. Por consequência, o modelo pode se tornar mais instável, causando divergências no treinamento desde o início. Com intuito de resolver essa questão, partes distintas da função de perda são ponderadas. A penalidade mais alta é para predições de coordenadas e a mais baixa para predições da probabilidade de confiança quando nenhuma onda está presente. Assim, a função de perda torna-se,

$$f(\cdot) = \rho_{coord} \sum_{i=0}^K \sum_{j=0}^J [(t_i - \hat{t}_i)^2 + (s_i - \hat{s}_i)^2] + \rho_{vazio} \sum_{i=0}^K \sum_{j=0}^J (B_{ij} - \hat{B}_{ij})^2. \quad (3.5)$$

O *SQE* trata da mesma forma os erros em janelas grandes e pequenas. No entanto, a métrica deve enfatizar que pequenos desvios em janelas grandes importam menos do que nas pequenas. Uma solução parcial é calcular o logaritmo natural,  $\ln(\cdot)$ , da largura e altura da janela,

no lugar dessas medidas diretamente,

$$\rho_{coord} \sum_{i=0}^K \sum_{j=0}^J [(\ln l_i - \ln \hat{l}_i)^2 + (\ln a_i - \ln \hat{a}_i)^2]. \quad (3.6)$$

Destacam-se, portanto, dois pontos importantes sobre o treinamento do modelo. O primeiro é que são fornecidos pesos diferentes para predições de confiança sobre as janelas que contêm ou não ondas durante o treinamento. A outra questão é o fato de se classificar o logaritmo natural da largura e altura de janelas para penalizar o erro em ondas pequenas grandes de maneira diferente.

O modelo prediz várias janelas por região. No momento do treinamento, espera-se que apenas uma classificação de janela seja responsável por cada onda. Dessa forma, atribui-se um preditor “pivô” por classificar uma onda com maior  $j(O_{real}, O_{predita})$ . Essa estratégia aumenta a especialização entre as predições das janelas das ondas. Cada pivô aprimora a classificação de determinados tamanhos, proporções ou tipos de batimentos, melhorando a sensibilidade do modelo.

A diferença entre o nível de confiança deve ser computada para cada classificação de janela.  $B_{ij}$  é a porcentagem de confiança esperada e  $\hat{B}_{ij}$  é o valor do modelo para a janela da onda predita,  $P(O) * j(O_{real}, O_{predita})$ ,

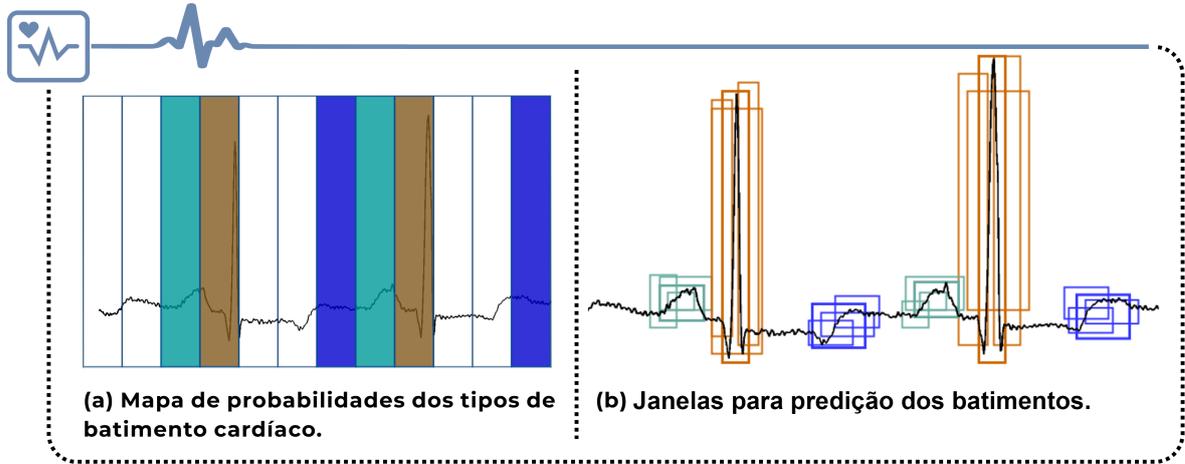
$$\sum_{i=0}^K \sum_{j=0}^J \mathbb{1}_{ij}^{onda} (B_{ij} - \hat{B}_{ij})^2 + \rho_{vazio} \sum_{i=0}^K \sum_{j=0}^J \mathbb{1}_{ij}^{vazio} (B_{ij} - \hat{B}_{ij})^2, \quad (3.7)$$

em que  $\mathbb{1}_{ij}^{onda}$  indica se a onda existe na  $i$ -ésima região e que a  $j$ -ésima janela é responsável por detectá-la. Por outro lado,  $\mathbb{1}_{ij}^{vazio}$  é 1 se não houver onda na  $i$ -ésima região.

Por fim, o fator para classificação usa a *SQE* para comparar as probabilidades condicionais entre todas as classes. Para simplificar o aprendizado, a função de perda deve considerar apenas erros de classificação das ondas que estão presente na região. Assim, defini-se a máscara  $\mathbb{1}_i^{onda}$ . A Figura 24 apresenta um exemplo do mapa de mapa de probabilidades dos tipos de batimento cardíaco e das  $J$  janelas criadas por região.

$$\sum_{i=0}^K \mathbb{1}_i^{onda} \sum_{c \in classes} (P_i(b) - \hat{P}_i(b))^2. \quad (3.8)$$

Figura 24 – Duas formas de visualizar a classificação dos batimentos cardíacos. (a) exemplo do mapa de probabilidades para os tipos de batimento cardíaco e (b)  $J$  janelas criadas por regiões.



Fonte: elaborado pelo autor.

A Equação 3.9 mostra a versão completa da função de perda adotada na otimização do problema multiclasse durante o treinamento.

$$\begin{aligned}
 f(\cdot) = & \rho_{coord} \sum_{i=0}^K \sum_{j=0}^J \mathbb{1}_i^{onda} [(t_i - \hat{t}_i)^2 + (s_i - \hat{s}_i)^2] \\
 & + \rho_{coord} \sum_{i=0}^K \sum_{j=0}^J [(\ln l_i - \ln \hat{l}_i)^2 + (\ln a_i - \ln \hat{a}_i)^2] \\
 & + \sum_{i=0}^K \sum_{j=0}^J \mathbb{1}_{ij}^{onda} (B_{ij} - \hat{B}_{ij})^2 + \rho_{vazio} \sum_{i=0}^K \sum_{j=0}^J \mathbb{1}_{ij}^{vazio} (B_{ij} - \hat{B}_{ij})^2 \\
 & + \sum_{i=0}^K \mathbb{1}_i^{onda} \sum_{c \in classes} (P_i(b) - \hat{P}_i(b))^2.
 \end{aligned} \tag{3.9}$$

Observa-se que a função de perda,  $f(\cdot)$ , só penaliza o erro de classificação se uma onda estiver presente nessa região, por isso a probabilidade condicional da classe foi apresentada anteriormente. Ela também só penaliza o erro de posição da janela se a classificação é referente ao pivô responsável pelo padrão ouro da janela. Isto é, tem o  $j(O_{real}, O_{predita})$  maior que qualquer classificação na região.

Alguns desafios podem ser destacados sobre o método. Como o modelo aprende janelas das ondas a partir de dados, ele se esforça para generalizar padrões de sinais em proporções incomuns. Outra questão é quanto a função de perda. Ela trata os erros em janelas pequenas da mesma forma que em grandes. Um pequeno erro em uma janela grande geralmente é benigno, mas um pequeno erro em uma janela pequena tem um efeito muito maior no  $j(O_{real}, O_{predita})$ . Dessa forma, os principais tipos de erros são localizações incorretas durante a detecção.

### 3.2.2.2 Junção das bases de segmentação e classificação durante o treinamento

Um outro desafio para treinar o método proposto para as tarefas simultâneas de segmentação das ondas e classificação das arritmias cardíacas consiste no fato da maioria das bases presentes na literatura tratar os problemas de maneira distinta. Dessa forma, alguns ajustes foram realizados para formação da base de treinamento.

Primeiramente, o método foi treinado apenas com a base de dados QT, cujo objetivo é segmentar as ondas P, T e QRS. O conhecimento aprendido nessa base foi utilizado para aprimorar o treinamento da base de dados de classificação e também para determinar as marcações do QRS na base MIT-BIH, uma vez que a mesma só possui a marcação do centro do QRS. Ou seja, utilizou-se o conceito de Transferência de Aprendizado (*Transfer Learning*) para o modelo final (TORREY; SHAVLIK, 2010).

Após o treinamento com a base QT e a transferência de conhecimento para MIT-BIH e posterior treinamento com a mesma, tivemos um modelo final com as 3 classes da QT (QRS, P e T), as quatro classes que indicam os tipos de arritmia cardíaca sobre o QRS na MIT-BIH e a classe QRS normal da última base. Além disso, a saída do modelo também foi capaz de obter as informações das janelas para as ondas com o mesmo treinamento.

### 3.2.3 Pós-processamento

Como pode ser observado na Figura 24(b), algumas janelas redundantes são criadas pelo método. O primeiro passo, naturalmente, é remover todas as janelas com baixa probabilidade. Embora essa passo elimine detecções anômalas de ondas, muitas janelas ainda seriam associadas ao mesmo batimento. Entretanto, cada onda deve ter apenas uma janela. Uma abordagem mais inteligente de seleção pode ser realizada considerando também as sobreposições das janelas por meio da métrica  $j(\cdot, \cdot)$ .

A abordagem de seleção de janela é apresentada no Algoritmo 1. O algoritmo começa com a escolha da janela com o maior nível de confiança, linha 4. Essa janela será adicionada em  $\mathcal{F}$ , que inicia vazia, e removida da lista inicial de predições (linhas 5 e 6). Em seguida, computa-se o  $j(\mathcal{M}, j_i)$  entre  $\mathcal{M}$  e todas as outras predições de janela, linhas 7-9. Se a sobreposição for maior que o limiar  $\gamma$ , a janela será removida da lista inicial  $\mathcal{J}$ . Esse processo é repetido para cada tipo de batimento até que não haja mais janelas em  $\mathcal{J}$ , linha 2. Essa técnica também é conhecida como supressão não máxima na literatura (LECUN *et al.*, 2015).

---

**Algoritmo 1** Algoritmo de seleção das janelas previstas pelo modelo.
 

---

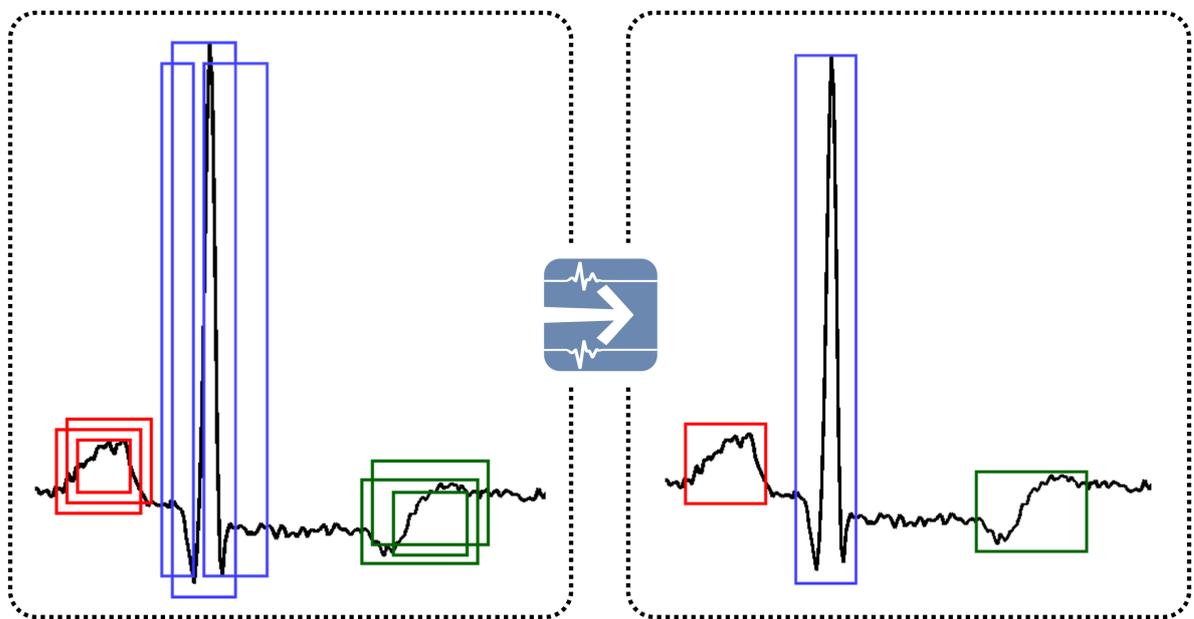
```

SELECAO-JANELA( $\mathcal{J}$ )
1   $\mathcal{F} = \emptyset$ 
2  Enquanto  $\mathcal{J} \neq \emptyset$ 
3       $m = \arg \max(J.prob)$ 
4       $\mathcal{M} = j_m$ 
5       $\mathcal{F} = \mathcal{F} \cup \mathcal{M}$ 
6       $\mathcal{J} = \mathcal{J} - \mathcal{M}$ 
7      Para  $j_i \in \mathcal{J}$ 
8          Se  $j(\mathcal{M}, j_i) > \gamma$ 
9               $\mathcal{J} = \mathcal{J} - j_i$ 
10 Retorna  $\mathcal{F}$ 
  
```

---

A Figura 25 exemplifica o resultado da abordagem de seleção de janelas utilizando a métrica de sobreposições das janelas,  $j(\cdot, \cdot)$ , e o nível de confiança. Inicialmente, o modelo prediz mais de uma janela para diferentes batimentos. Após o algoritmo de seleção, apenas uma janela é responsável por caracterizar o batimento cardíaco.

Figura 25 – Predições antes e depois do algoritmo de seleção de janela sobre cada tipo de batimento.



Fonte: elaborado pelo autor.

### 3.3 Especificações para caracterização das arritmias cardíacas no eletrocardiograma

Com objetivo de atender aos padrões clínicos, o desempenho do método proposto nesta tese no problema de classificação de arritmias cardíacas foi avaliado seguindo as recomendações ANSI/AAMI EC57 da *Advancement of Medical Instrumentation* (AAMI). Uma vez que o método também detecta e segmenta os batimentos cardíacos concomitantemente à classificação, também verificou-se o comportamento nessas duas etapas. Além disso, os resultados do método proposto são confrontados com de outros trabalhos estado da arte.

#### 3.3.1 Descrição da base MIT-BIH

No grupo de experimentos da classificação, utilizou-se a base de dados de arritmias cardíacas MIT-BIH (Massachusetts Institute of Technology - Boston Beth Israel Hospital) *Arrhythmia Database* (MOODY; MARK, 2001) disponibilizada online por (GOLDBERGER *et al.*, 2000) e recomendada pela norma da ANSI/AAMI EC57. Essa base, a qual será referenciada na tese como MIT-BIH, é um conjunto de sinais de eletrocardiograma amplamente adotada para avaliação do desempenho de algoritmos para a classificação de arritmias (LUZ *et al.*, 2016). Os dados consistem em 48 trechos de 30 minutos de registro de sinais de ECG, provenientes de gravações de 24 horas de duração.

A fonte dos ECGs incluídos no banco de dados de arritmia do MIT-BIH é um conjunto de mais de 4.000 registros de *Holter* de longo prazo obtidos pelo Laboratório de arritmia do Hospital Beth Israel entre 1975 e 1979. Aproximadamente 60% desses registros foram obtidos de pacientes internados. O banco de dados contém 23 registros escolhidos aleatoriamente a partir deste conjunto, e 25 registros selecionados do mesmo conjunto para incluir uma variedade de fenômenos raros, mas clinicamente importantes, que não seriam bem representados por uma pequena amostra aleatória de gravações de *Holter*.

O primeiro grupo destina-se a servir como uma amostra representativa da variedade de formas de onda e artefatos que um detector de arritmia pode encontrar no uso clínico de rotina. Por outro lado, o segundo grupo foi escolhido para incluir arritmias ventriculares, juncionais e supraventriculares complexas e anormalidades de condução. Os pacientes foram 25 homens com idades entre 32 a 89 anos e 22 mulheres com idades entre 23 e 89 anos.

Na maioria dos registros, o sinal superior é uma derivação modificada do membro II (MLII), obtida colocando os eletrodos no tórax. O sinal mais baixo é geralmente uma derivação

modificada V1; quanto ao sinal superior, os eletrodos também são colocados no tórax. Esta configuração é usada rotineiramente pelo Laboratório de Arritmia BIH. Os complexos QRS normais são geralmente proeminentes no sinal superior. O eixo principal para o sinal inferior pode ser quase ortogonal ao eixo elétrico cardíaco médio, no entanto.

Os sinais de ECG foram amostrados em dois canais. O primeiro foi obtido por meio de eletrodos colocados no peito do paciente, em uma derivação modificada DII localizada entre o braço direito e a perna esquerda. Enquanto, no outro canal, os eletrodos foram posicionados no ponto pericardial V1, podendo ter sido captado também nos pontos V2, V4 ou V5, dependendo do paciente.

Os sinais foram digitalizados a partir das fitas *Holter* a uma taxa de amostragem de 360Hz. Os batimentos foram marcados e classificados manualmente por especialistas em 15 classes, quanto ao tipo de arritmia.

Há quatro registros provenientes de pacientes que utilizavam marca-passo na MIT-BIH. Estes registros não foram considerados por recomendação da norma ANSI/AAMI EC57. Com esta exclusão, os 44 registros considerados não possuem batimentos dos tipos / e f.

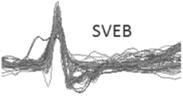
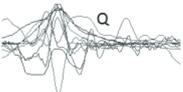
### **3.3.2 Agrupamento dos batimentos cardíacos e atributos dos extratores de características**

A norma ANSI/AAMI EC57 recomenda agrupar as anotações em cinco classes, conforme mostra a Tabela 2. Além dessa divisão, o método proposto nesta tese também é avaliado com o agrupamento sugerido por Llamedo e Martínez (2010), que divide o problema em três classes principais: N, SVEB e VEB. Essa redução no grupo de classes é motivada principalmente pela quantidade não representativa de amostras da classe Q e os valores inexpressivos alcançados pela F nos trabalhos relatados na literatura. Avaliou-se também o desempenho do método em uma abordagem binária que agrupa os batimentos nas classes normal ou com arritmia cardíaca.

A norma AAMI também recomenda dividir os dados em dois conjuntos de dados para treinamento e validação dos modelos, mas sem especificar o uso de quais registros ou batimento cardíaco. Diversos trabalhos na literatura não separam os batimentos cardíacos do mesmo paciente para formar esses dois conjuntos. Esta abordagem é conhecida como paradigma intra-paciente (LUZ *et al.*, 2016).

Em um ambiente clínico, no entanto, um método totalmente automático encontrará os batimentos cardíacos dos pacientes diferentes daqueles que eles costumavam aprender na fase de treinamento, denominado paradigma inter-paciente. Chazal *et al.* (2004) propôs dividir

Tabela 2 – Agrupamento das anotações da MIT-BIH segundo a norma ANSI/AAMI EC57.

Classe AAMI	Anotação médica	Tipo de Batimento	Forma da onda
Normal (N)	N	Batimento normal	
	L	Bloqueio de ramo esquerdo	
	R	Bloqueio de ramo direito	
	e	Escape atrial	
	j	Escape nodal	
Batimento supraventricular ectópico (SVEB)	A	Batimento atrial prematuro	
	a	Batimento atrial prematuro aberrado	
	j	Batimento prematuro nodal	
	S	Batimento supraventricular prematuro	
Batimento ventricular ectópico (VEB)	V	Contração ventricular prematura	
	E	Escape ventricular	
Batimento de fusão (F)	F	Fusão de batimento ventricular e normal	
Batimento desconhecido (Q)	P	Batimento de marca-passo	
	f	Fusão de batimentos normal e de marca-passo	
	U	Batimento inclassificável	

Fonte: elaborado pelo autor.

o batimento cardíaco do banco de dados MIT-BIH em dois conjuntos de dados, DS1 e DS2, com registros de indivíduos diferentes entre eles como mostra a Tabela 3. Além desses dois paradigmas, também avaliou-se os resultados do método proposto sobre todo o conjunto de dados.

Tabela 3 – Conjuntos de dados propostos por Chazal *et al.* (2004) para treinar (DS1) e avaliar os modelos (DS2).

Divisão	Registros
DS1	101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 220, 223, 230
DS2	100, 103, 105, 11, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233, 234

Fonte: elaborado pelo autor.

As Tabelas 4 e 5 mostram a quantidade de atributos dos extratores de características usados na abordagem clássica de classificação de arritmia cardíaca para a divisão com 5 e 3 classes considerando a divisão do DS1 e DS2.

### 3.4 Base de dados para segmentação e detecção das ondas no ECG

Visto que o método proposto também é capaz de detectar e segmentar os batimentos cardíacos, avaliou-se seu desempenho nas tarefas supracitadas com a base de dados QT (LA-

Tabela 4 – Quantidade de atributos considerando a divisão com 5 classes.

Divisão	Extrator de Característica	$n_{atr}$	N	SVEB	VEB	F	Q	Total	
DS1	<b>Fourier</b>	33	45844	943	3788	415	8	50998	
	<b>Goertzel</b>	33	45844	943	3788	415	8	50998	
	<b>HOS</b>	4	45844	943	3788	415	8	50998	
	<b>SCM</b>	8	45844	943	3788	415	8	50998	
	<b>SCM-Fourier</b>	8	45844	943	3788	415	8	50998	
	Chazal <i>et al.</i> (2004)	155	45747	940	3777	415	8	50887	
	Güler e Übeyli (2005)	19	45845	943	3788	415	8	50999	
	Song <i>et al.</i> (2005)	21	45825	943	3788	414	8	50978	
	Yu e Chen (2007)	13	45844	943	3788	415	8	50998	
	Yu e Chou (2008)	31	45511	929	3770	412	8	50630	
	Ye <i>et al.</i> (2010)	100	45844	943	3788	415	8	50998	
	DS2	<b>Fourier</b>	33	44238	1836	3221	388	7	49690
		<b>Goertzel</b>	33	44238	1836	3221	388	7	49690
<b>HOS</b>		4	44238	1836	3221	388	7	49690	
<b>SCM</b>		8	44238	1836	3221	388	7	49690	
<b>SCM-Fourier</b>		8	44238	1836	3221	388	7	49690	
Chazal <i>et al.</i> (2004)		155	44181	1786	3218	388	7	49580	
Güler e Übeyli (2005)		19	44238	1836	3221	388	7	49690	
Song <i>et al.</i> (2005)		21	44218	1836	3219	388	7	49668	
Yu e Chen (2007)		13	44238	1836	3221	388	7	49690	
Yu e Chou (2008)		31	43905	1823	3197	388	7	49320	
Ye <i>et al.</i> (2010)		100	44238	1836	3221	388	7	49690	

Fonte: elaborado pelo autor.

Tabela 5 – Quantidade de atributos considerando a divisão com 3 classes.

Divisão	Extrator de Característica	$n_{atr}$	N	SVEB	VEB	Total	
DS1	<b>Fourier</b>	33	45844	943	4211	50998	
	<b>Goertzel</b>	33	45844	943	4211	50998	
	<b>HOS</b>	4	45844	943	4211	50998	
	<b>SCM</b>	8	45844	943	4211	50998	
	<b>SCM-Fourier</b>	8	45844	943	4211	50998	
	Chazal <i>et al.</i> (2004)	155	45747	940	4200	50887	
	Güler e Übeyli (2005)	19	45845	943	4211	50999	
	Song <i>et al.</i> (2005)	21	45825	943	4210	50978	
	Yu e Chen (2007)	13	45844	943	4211	50998	
	Yu e Chou (2008)	31	45511	929	4190	50630	
	Ye <i>et al.</i> (2010)	100	45844	943	4211	50998	
	DS2	<b>Fourier</b>	33	44238	1836	3616	49690
		<b>Goertzel</b>	33	44238	1836	3616	49690
<b>HOS</b>		4	44238	1836	3616	49690	
<b>SCM</b>		8	44238	1836	3616	49690	
<b>SCM-Fourier</b>		8	44238	1836	3616	49690	
Chazal <i>et al.</i> (2004)		155	44181	1786	3613	49580	
Güler e Übeyli (2005)		19	44238	1836	3616	49690	
Song <i>et al.</i> (2005)		21	44218	1836	3614	49668	
Yu e Chen (2007)		13	44238	1836	3616	49690	
Yu e Chou (2008)		31	43905	1823	3592	49320	
Ye <i>et al.</i> (2010)		100	44238	1836	3616	49690	

Fonte: elaborado pelo autor.

GUNA *et al.*, 1997). Esta base possui 105 sinais ECG, com duração de 15 minutos cada. Ela inclui registro das bases MIT-BIH além de diversas outras do sistema *Physionet* (*European ST-T database, ST change, Supraventricular Arrhythmia, Normal Sinus Rhythm, Sudden Death e Long Term*).

A elaboração da base QT tem como princípio a validação de métodos de segmentação das ondas características do eletrocardiograma. As anotações dentro de cada registro, entre 30 e 100 batimentos representativos foram anotados manualmente por cardiologistas, que identificaram os picos (pontos fiduciais) dos complexos QRS e das ondas P e T, além dos inícios e fins dessas ondas. Os sinais foram digitalizados a uma taxa de amostragem de 250 Hz.

### 3.5 Métricas de avaliação

Com objetivo de analisar o desempenho do método proposto na tarefa de classificação de arritmias cardíacas, cinco métricas de avaliação são empregadas e definidas a seguir: acurácia, sensibilidade, especificidade, valor preditivo positivo, *F1-Score* e média harmônica. Essas medidas são obtidas a partir da matriz de confusão. O cálculo de cada categoria é dado da seguinte forma:

- Verdadeiro Positivo (VP): é o número de classificações corretas de batimentos cardíacos sem arritmia.
- Falso Negativo (FN): corresponde a quantidade de batimentos cardíacos normais classificado erroneamente como arritmica.
- Falso Positivo (FP): indica os batimentos cardíacos arrítmicos classificados incorretamente como normais.
- Verdadeiro Negativo (VN): corresponde ao número de classificações corretas de batimentos cardíacos do paciente com arritmia.

A Figura 26 mostra um exemplo da matriz de confusão para a versão de 5 classes, enquanto a Tabela 6 sumariza as métricas adotadas. Com objetivo de facilitar a compreensão, destacou-se a classe normal em negrito, uma vez que o VP, FN, FP e VN são relativos às classes analisadas.

Quanto ao desempenho da detecção de batimentos cardíacos, de acordo com Kohler *et al.* (2002), as métricas de avaliação essenciais são a sensibilidade (*Se*) e o valor preditivo

Figura 26 – Exemplo da matriz de confusão para um problema multiclasse.

		Classe Predita		
		s, v	n	f, q
Classe Real	f, q	VN	FP	VN
	n	FN	VP	FN
	s, v	VN	FP	VN

Fonte: elaborado pelo autor.

Tabela 6 – Resumo das métricas adotadas na classificação das arritmias cardíacas.

Métrica	Equação	Descrição
Acurácia (Ac)	$\frac{VP + VN}{Total}$	No geral, com que frequência o método está correto?
Sensibilidade (Se)	$\frac{VP}{VP + FN}$	Quando na verdade é arritmia, com que frequência o método prediz como arritmia?
Especificidade (Es)	$\frac{VN}{VN + FP}$	Quando é realmente normal, com que frequência o método prediz normal?
Valor Preditivo Positivo (VPP)	$\frac{VP}{VP + FP}$	Quando prediz arritmia, com que frequência é arritmia?
F1-Score (FS)	$2 \cdot \frac{Se \cdot VPP}{Se + VPP}$	É a média ponderada entre Sensibilidade e o Valor Preditivo Positivo.
Média Harmônica (MH)	$2 \cdot \frac{Se \cdot Es}{Se + Es}$	É a média ponderada entre Sensibilidade e a Especificidade.

Fonte: elaborado pelo autor.

positivo ( $VPP$ ), definidas por

$$Se = \frac{TP}{TP + FN}, \quad (3.10)$$

$$VPP = \frac{TP}{TP + FP}.$$

em que TP denota a quantidade de detecções corretas, FN indica o número de vezes que um batimento verdadeiro não foi detectado e FP total de ocorrências que o modelo detectou um batimento não verdadeiro, ou seja, outro artefato.

No estudo da segmentação das ondas, a diferença temporal entre as ondas segmentadas pelo método proposto e as marcações dos especialistas presente na base QT foi avaliada. Além disso, verificou-se a diferença entre os limites segmentados na matriz de transformação

**M** do método proposto. Essa diferença foi normalizada pelo tamanho do sinal predito, que é o mesmo da matriz.

Além das métricas supracitadas, computou-se o custo computacional do método proposto para treinar e classificar um registro de ECG. Os experimentos foram realizados no Google Colaboratory (PC Intel Xeon rodando a 2,30 GHz, 13 GB de RAM e NVIDIA Tesla K80 com sistema operacional Linux) e codificados na linguagem Python.

Os parâmetros necessários durante a análise do método proposto para detectar, segmentar e classificar arritmias cardíacas são apresentados na Tabela 7. Eles foram fixados por meios testes experimentais exaustivos e independem das bases de dados.

Tabela 7 – Conjunto de parâmetros usado pelo método proposto nesta tese.

Parâmetro	Valor	Significado
$K$	7	Número de regiões.
$J$	2	Número de janelas.
$\rho_{coord}$	5	Penalidade para classificações das coordenadas.
$\rho_{vazio}$	0,5	Penalidade para classificações da probabilidade de confiança quando nenhuma onda está presente.
$\gamma$	95%	Limiar usado para seleção das janelas.

Fonte: elaborado pelo autor.

### 3.6 Considerações finais

Neste capítulo, foi apresentado a metodologia empregada na pesquisa. Inicialmente, buscou-se a melhor combinação de extratores de características e classificadores que usam a abordagem clássica para classificação de arritmias cardíacas em sinais de ECG.

As dificuldades provenientes da primeira etapa de estudo motivaram a criação de um novo método para tarefa de interesse que usa Aprendizado Profundo. Ao contrário das abordagens convencionais de processamento digital de sinais que necessita de técnicas distintas nas etapas para classificação da anomalia cardíaca, o modelo proposto realiza a tarefa de ponta a ponta.

Em seguida, as especificações aderidas para caracterização das arritmias cardíacas à luz da norma ANSI/AAMI EC57 são explanadas. Além disso, o capítulo mostrou a base de dados adotada para avaliar a eficiência do método proposto para segmentação e detectar as ondas no ECG. Por fim, as métricas para avaliação são apresentadas.

## 4 RESULTADOS

Os resultados desta tese são apresentados em dois grandes grupos de experimentos. Inicialmente, avaliou-se diferentes técnicas de extração de características e classificação no problema de classificação de anomalias cardíacas utilizando sinais de eletrocardiograma. Em seguida, verificou-se a eficiência do método que através de uma abordagem de regressão com Aprendizado Profundo é capaz de simultaneamente prever arritmias cardíacas, detectar e segmentar os batimentos em sinais de eletrocardiograma. É importante ressaltar que um único modelo foi otimizado para realizar simultaneamente estas tarefas no último grupo de estudo.

### 4.1 Resultados da classificação de arritmia cardíaca com a abordagem clássica

Esta seção apresenta e discute os resultados da abordagem clássica para classificação de anomalias cardíacas em ECG. Primeiramente, são apresentados os melhores resultados com os paradigmas intra-paciente e inter-paciente da literatura. Posteriormente, os resultados da abordagem clássica desta tese são comparados com os extratores que obtiveram melhor desempenho.

Os experimentos foram projetados para responder à pergunta: **Qual a combinação de extrator de características e classificador é mais adequada para classificação de arritmia cardíaca?** As discussões também levam em consideração aqueles que obtiveram acurácia superior ou equivalente aos trabalhos do estado da arte, bem como o tempo necessário para a obtenção dos resultados.

Luz *et al.* (2016) fizeram uma extensa revisão bibliográfica que agrupa as pesquisas para classificação de arritmias cardíacas usando sinais ECG com base no sistema de avaliação utilizado. A Tabela 9 apresenta os trabalhos que seguem o paradigma intra-paciente, onde os batimentos cardíacos dos mesmos pacientes provavelmente aparecem tanto no treinamento quanto no conjunto de dados de teste. Enquanto a Tabela 8 mostra os trabalhos que seguem o paradigma inter-paciente, onde os autores tiveram o cuidado ao construir e avaliar a classificação utilizando os batimentos cardíacos de diferentes pacientes, seguindo o protocolo proposto por Chazal *et al.* (2004).

Os trabalhos que não se enquadram no paradigma inter-paciente, não permitem comparações justas com os resultados da literatura, uma vez que a grande maioria dos autores não seguiu o mesmo protocolo para as avaliações. Como pode ser observado na Tabela 9, também

Tabela 8 – Métodos da abordagem clássica que usam o paradigma inter-paciente.

Trabalho	Extrator	Classificador	Resultados
Esta Tese	HOS; SCM; Fourier; Goertzel	Bayes; MLP; SVM; OPF	Ac: 94%; $Se_N$ : 99%; $Es_N$ : 78%; $Se_S$ : 95%; $Es_S$ : 91%; $Se_V$ : 88%; $Es_S$ : 99%
Chazal <i>et al.</i> (2004)	Intervalo ECG; morfologia	LD	Ac: 83%; $Se_N$ : 87%; $VPP_N$ : 99%; $Se_S$ : 76%; $VPP_S$ : 38%; $Se_V$ : 77%; $VPP_V$ : 82%
Soria e Martínez (2009)	Intervalo RR; VCG; morfologia + FFS	LD	Ac: 90%; $Se_N$ : 92%; $VPP_N$ : 85%; $Se_S$ : 88%; $VPP_S$ : 93%; $Se_V$ : 90%; $VPP_V$ : 92%
Llamedo e Martínez (2010)	Wavelet; VCG + SFFS	LD	Ac: 93%; $Se_N$ : 95%; $VPP_N$ : 98%; $Se_S$ : 77%; $VPP_S$ : 39%; $Se_V$ : 81%; $VPP_V$ : 87%
Mar <i>et al.</i> (2011)	Morfologia e tempo; atributos estatísticos + SFFS	LD; MLP	Ac: 89%; $Se_N$ : 89%; $VPP_N$ : 99%; $Se_S$ : 83%; $VPP_S$ : 33%; $Se_V$ : 86%; $VPP_V$ : 75%
# Bazi <i>et al.</i> (2013)	Morfologia; Wavelet	SVM; DTSVM	Ac: 97% (DS1); Ac: 92% (DS2)
Luz <i>et al.</i> (2013)	Wavelet; ICA; PCA; intervalo RR	SVM; MLP; Bayes; OPF	$Se_N$ : 84%; $Es_{SVEB}$ : 18%; $Es_{VEB}$ : 72%
Ye <i>et al.</i> (2012)	Morfologia; Wavelet; Intervalo RR; ICA; PCA	SVM	Ac: 86.4%; $Se_N$ : 88%; $VPP_N$ : 97%; $Se_S$ : 60%; $VPP_S$ : 53%; $Se_V$ : 81%; $VPP_V$ : 63%
Lannoy <i>et al.</i> (2010)	Intervalo ECG; morfologia; HOS; coeficientes HBF	SVM	Ac: 83%; $Se_N$ : 80%; $Se_S$ : 88%; $Se_V$ : 78%
Park <i>et al.</i> (2008)	HOS; HBF	SVM hierárquico	Ac: 85%; $Se_N$ : 86%; $Se_S$ : 82%; $Se_V$ : 80%
Zhang e Luo (2014)	Intervalo RR; morfologia; intervalo ECG e segmentos	SVM	Ac: 86%; $Se_N$ : 89%; $VPP_N$ : 99%; $Se_S$ : 79%; $VPP_S$ : 35%; $Se_V$ : 85%; $VPP_V$ : 92%;
# Huang <i>et al.</i> (2014)	Projeção aleatória; intervalo RR	SVM	$Se_N$ : 99%; $VPP_N$ : 95%; $Se_S$ : 91%; $VPP_S$ : 42%; $Se_V$ : 94%; $VPP_V$ : 91%
\$ Lin e Yang (2014)	intervalo RR	LD	Ac: 93%; $Se_N$ : 91%; $VPP_N$ : 99%; $Se_S$ : 81%; $VPP_S$ : 31%; $Se_V$ : 86%; $VPP_V$ : 73%
Lannoy <i>et al.</i> (2011)	Intervalo RR; morfologia dos segmentos ECG; HBF; HOS	CRF	Ac: 85%; $Ac_N$ : 79%; $Ac_S$ : 92%; $Ac_V$ : 85%
Zhang e Luo (2014)	Intervalo RR; morfologia; intervalo ECG e segmentos; Wavelet	SVM	Ac: 87%; $Se_N$ : 88%; $VPP_N$ : 98%; $Se_S$ : 74%; $VPP_S$ : 59%; $Se_V$ : 88%; $VPP_V$ : 82%

CRF: *Conditional Random Fields*, FFS: *Floating Feature Selection*, HBF: *Hermite Basis Function*, HOS: *Higher-Order Statistics*, ICA: *Independent Component Analysis*, LD: *Linear Discriminants*, MLP: *Multilayer Perceptron*, PCA: *Principal Component Analysis*, SCM: *Structural Co-occurrence Matrix*, SFFS, *Sequential forward floating search*, VCG: *Vetorcardiograma*. \$ Os autores otimizam seus resultados para 3 classes. # A matriz de confusão não foi fornecida, alguns valores não puderam ser calculados.

Fonte: Adaptado de Luz *et al.* (2016).

é difícil avaliar qual técnica contribui para a classificação dos batimentos cardíacos, uma vez que métodos com diferentes abordagens alcançam acurácias acima de 98%. Assim, os resultados reportados na Tabela 9 não podem ser considerados do ponto de vista clínico, uma vez que os valores reportados por estes trabalhos são provavelmente diferentes no cenário real em termos de acurácia.

Tabela 9 – Métodos da abordagem clássica que usam o paradigma intra-paciente.

Trabalho	# Cl.	Extrator	Classificador	Resultados
Chen <i>et al.</i> (1996)	2	Intervalo RR	Conjunto de regras	Ac: 95%
Lagerholm <i>et al.</i> (2000)	16	HBF, SOM	Clusterização	Ac: 98%
Dokur e Ölmez (2001)	10	Fourier, Wavelet + FSDP	MLP, RCE	Ac: 96%
Oowski e Linh (2001)	6	HOSC	NN fuzzy	Ac: 96%
Engin (2004)	4	HOSC, Wavelet	Dist. mín., kNN, Bayes	Ac: 98%
Güler e Übeyli (2005)	4	Wavelets	NN	Ac: 96%
Song <i>et al.</i> (2005)	6	Wavelet, LDA, Intervalo RR	SVM	Ac: 99%
Karimifard <i>et al.</i> (2006)	7	HBF	kNN	Ac: 99%
Bortolan <i>et al.</i> (2007)	2	VCG e morfologia, GA	Agrupamento Fuzzy	Ac: 99%
Yu e Chen (2007)	5	ICA, intervalo RR	PNN	Ac: 99%
Korürek e Nizam (2008)	6	Intervalo RR, segmentos ECG	Cluster baseado em ACO, kNN	Ac: 94%
Yu e Chou (2008)	8	Intervalo RR, ICA	PNN, BPNN	Ac: 98%
Kim <i>et al.</i> (2009)	6	Intervalo RR, PCA	ELM	Ac: 98%
Ye <i>et al.</i> (2010)	15	Wavelet, ICA, PCA, intervalo RR	SVM	Ac: 99%
Mishra e Raghav (2010)	6	LFD	kNN	Ac: 89%
Lanata <i>et al.</i> (2011)	6	HOS	MOG, kNN	Ac: 85%
Kallas <i>et al.</i> (2012)	3	KPCA	SVM	Ac: 97%
Tran <i>et al.</i> (2014)	7	Intervalo RR, HBF	Comitê de classificadores	Ac: 98%
Alickovic e Subasi (2015)	5	Modelagem autoregressiva (AR)	SVM, MLP, RBF, kNN	Ac: 99%

ACO: *Ant Colony Optimization based clustering*, BPNN: *Back Propagation Neural Network*, ELM: *Extreme Learning Machine*, HBF: *Hermite Basis Function*, HMM: *Hidden Markov modeling*, HOS: *Higher order statistics*, ICA: *Independent Component Analysis*, LDA: *Linear Discriminant Analysis*, LFD: *Local Fractal Dimension*, MOG: *Mixture of Gaussian*, NN: *Neural Network*, PCA: *Principal Component Analysis*, PNN: *Probabilistic neural Network*, SOM: *Self-Organizing Maps*.

Fonte: Adaptado de Luz *et al.* (2016).

Verifica-se, portanto, que a maioria dos trabalhos na literatura não se preocupa em seguir a divisão definida em (CHAZAL *et al.*, 2004) ou qualquer outro protocolo inter-paciente que impõe o não uso de batimentos cardíacos do mesmo paciente no treinamento e teste, conforme mostrado na Tabela 9.

Luz *et al.* (2013) reavaliaram alguns extratores de características que apresentavam uma acurácia próxima de 100%, mas não seguiam nenhum protocolo de seleção de batimentos cardíacos. Posteriormente, os autores reavaliaram os resultados produzidos pelos métodos com o objetivo de relatar os experimentos de acordo com o protocolo recomendado pela AAMI e utilizaram o esquema de divisão proposto por Chazal *et al.* (2004). A próxima seção compara e discute os resultados da abordagem clássica desta tese com o trabalho de Luz *et al.* (2013).

#### **4.1.1 Resultado dos métodos clássicos utilizando o protocolo inter-paciente**

A Tabela 10 apresenta os resultados da combinação entre os extratores e os classificadores para a divisão de 5 classes. Inicia-se com os ganhos obtidos pela abordagem com os métodos clássicos proposta nesta tese, enquanto a parte inferior mostra os melhores resultados reportados por Luz *et al.* (2013). As combinações com maiores acurácias são destacadas em verde. Os resultados da primeira parte da tabela são valores médios provenientes das diferentes realizações da tese, os outros trabalhos na literatura não deixam claro se fizeram mais de uma realização.

Comparando os melhores resultados com Luz *et al.* (2013): a combinação do HOS e Bayes alcançou 94,3% de acurácia enquanto o melhor resultado de Luz *et al.* (2013) foi com SVM, alcançando 92,2%, com as características propostas por Chazal *et al.* (2004). Destaca-se também os resultados para classificação de contração ventricular prematura (classe V), cuja abordagem proposta nesta tese obteve 87,8% de sensibilidade em comparação com 48% por Luz *et al.* (2013). Os resultados para as classes N, S, F e Q são semelhantes.

O extrator SCM-Fourier teve melhor desempenho com o classificador Bayes, cuja acurácia foi de 92,4%. Mesmo usando a porcentagem da DC recomendada por Peixoto e Filho (2018) para obter as características mais relevantes, 1 a 5%, os resultados ainda ficaram abaixo do extrator HOS.

O HOS foi o extrator com maior acurácia, 94,3% usando o Bayes. A semelhança de um QRS normal com uma função de densidade de probabilidade Gaussiana, juntamente com sua forma de calcular as dissimilaridades entre um sinal de entrada e uma referência Gaussiana pura beneficiou o método.

Durante a busca para compreender o bom desempenho do HOS em relação aos atributos propostos por Chazal *et al.* (2004), verificou-se que os autores utilizaram atributos baseados em referências de tempo para medir a morfologia do batimento cardíaco no ECG.

Tabela 10 – Resultados das combinações de extratores e classificadores na base de dados MIT-BIH considerando a divisão de 5 classes com a abordagem clássica.

Classificador	Ac	N		S		V		F		Q						
		Se	Es	MH	Se	Es	MH	Se	Es	MH	Se	Es	MH			
<b>Fourier</b>																
Bayes	20,2	16,6	80,0	27,5	92,2	19,7	32,4	29,1	90,6	44,1	8,8	74,9	15,7	14,3	93,6	24,8
MLP	76,4	81,0	49,5	61,4	10,1	94,0	18,3	60,8	85,1	70,9	1,0	98,0	2,0	0,0	100,0	0,0
OPF	83,0	92,5	26,9	41,7	5,3	96,0	10,1	38,7	96,1	55,2	0,1	99,1	0,3	0,0	100,0	0,0
SVM	91,1	99,6	22,2	36,3	0,0	100,0	0,0	36,9	99,4	53,8	0,0	100,0	0,0	0,0	100,0	0,0
<b>HOS</b>																
Bayes	94,3	99,0	77,6	87,0	95,4	91,8	93,4	87,8	99,4	93,2	86,5	89,6	88,0	28,0	95,0	43,3
MLP	91,3	96,0	53,8	68,9	0,0	100,0	0,0	89,2	96,8	92,8	0,3	99,0	0,5	14,3	100,0	25,0
OPF	87,7	93,4	43,5	59,4	2,1	96,0	4,1	71,6	97,7	82,6	0,2	99,1	0,4	0,0	100,0	0,0
SVM	89,0	100,0	0,0	0,0	0,0	100,0	0,0	0,0	100,0	0,0	0,0	100,0	0,0	0,0	100,0	0,0
<b>Goertzel</b>																
Bayes	65,9	68,4	58,4	63,0	85,2	72,8	78,5	17,5	98,9	29,8	83,0	92,4	87,5	14,3	98,5	25,0
MLP	77,8	83,5	36,2	50,5	3,7	97,1	7,0	51,1	85,7	64,1	0,0	98,2	0,0	0,0	100,0	0,0
OPF	80,6	92,6	24,4	38,6	8,7	96,1	16,0	33,9	96,4	50,1	0,0	99,0	0,0	25,0	100,0	40,0
SVM	90,5	99,8	15,2	26,4	0,0	100,0	0,0	25,6	99,7	40,7	0,0	100,0	0,0	0,0	100,0	0,0
<b>SCM</b>																
Bayes	85,0	88,9	56,0	68,7	8,9	99,3	16,3	85,8	89,1	87,4	0,0	100,0	0,0	14,3	99,9	25,0
MLP	90,9	99,2	23,3	37,8	0,0	99,9	0,0	39,2	99,2	56,2	0,0	100,0	0,0	0,0	100,0	0,0
OPF	84,0	91,6	26,2	40,7	4,3	95,9	8,2	32,7	95,5	48,7	1,0	99,1	2,0	16,7	100,0	28,6
SVM	90,9	99,9	18,4	31,0	0,0	100,0	0,0	30,9	99,8	47,2	0,0	100,0	0,0	0,0	100,0	0,0
<b>SCM-Fourier</b>																
Bayes	92,4	97,5	52,5	68,3	18,8	98,8	31,6	74,8	99,2	85,3	9,3	99,3	17,0	0,0	100,0	0,0
MLP	91,7	99,0	33,6	50,2	0,6	99,5	1,2	56,0	99,4	71,7	0,0	99,9	0,0	0,0	100,0	0,0
OPF	86,9	92,3	38,2	54,0	3,6	96,0	6,9	62,3	96,4	75,7	1,3	99,1	2,5	0,0	100,0	0,0
SVM	89,7	100,0	6,1	11,6	0,0	100,0	0,0	10,3	99,9	18,7	0,0	100,0	0,0	0,0	100,0	0,0
<b>Extratores de características clássicos adotados em (LUZ et al., 2013)</b>																
<b>Chazal et al. (2004)</b>																
SVM	92,2	99,6	34,5	51,2	0,0	100,0	0,0	48,0	99,5	64,8	48,7	99,8	65,5	0	100,0	0
<b>Song et al. (2005)</b>																
SVM	89,3	100,0	3,0	05,7	0,0	100,0	0,0	4,8	100,0	9,2	0,0	100,0	0,0	0	100,0	0
<b>Yu e Chou (2008)</b>																
MLP	92,0	97,1	55,2	70,3	3,8	99,5	7,3	83,0	98,2	89,9	4,9	99,1	9,3	0	100,0	0
<b>Güler e Übeyli (2005)</b>																
SVM	89,4	99,9	4,8	92,0	0,0	100,0	0,0	8,0	99,9	14,8	0,0	100,0	0,0	0	100,0	0
<b>Yu e Chen (2007)</b>																
MLP	89,0	95,3	41,2	57,5	0,0	100,0	0,0	64,3	95,2	76,8	0,0	100,0	0,0	0	100,0	0
<b>Ye et al. (2010)</b>																
SVM	91,9	99,8	33,6	50,3	0,0	100,0	0,0	47,6	99,2	64,3	0,0	100,0	0,0	0	100,0	0

Fonte: elaborado pelo autor.

Chazal *et al.* (2004) desenvolveram um método que se ajusta ao conjunto de dados MIT-BIH, cujas configurações são únicas, como a frequência de amostragem. Por outro lado, a abordagem clássica proposta nesta tese foi pensada para ser mais genérica e buscou-se um método como o HOS que é intrínseco a uma distribuição de probabilidade do sinal. Além disso, o HOS possui apenas 4 atributos, ao contrário dos 155 propostos no trabalho de Chazal *et al.* (2004).

A SCM-Fourier com Bayes obteve acurácia de 92,4% e especificidade acima de 98% na maioria das classes. Par uso clínico, isso indica que a combinação tende a fornecer menos predições de falso positivo para pacientes saudáveis. Quanto a escolha do filtro na SCM, a SCM-Fourier com o sinal reconstruído com 1% da componente DC teve um ganho de 2,27% em relação à SCM com filtro da média.

Os resultados de Fourier e Goertzel foram inferiores aos dos outros extratores. Ambos os métodos se baseiam no domínio da frequência e um parâmetro chave é ter uma resolução boa o suficiente para identificar as harmônicas do sinal. A frequência de amostragem da base de dados, no entanto, diminui a resolução no espectro de frequência. Abordagens para resolver esse problema, no entanto, mudariam a metodologia proposta pelos outros trabalhos, tornando a comparação injusta.

Ao analisar os resultados da versão com 3 classes no conjunto de dados MIT-HIB, Tabela 11, verifica-se que o HOS teve desempenho semelhante a da abordagem com 5 classes. Esse extrator juntamente com o Bayes obtiveram acurácia de 94%, sensibilidade acima de 99% para os batimentos cardíacos normais e especificidade de 99% nas classes VEB e SVEB. Além disso, esses resultados são superiores aos do trabalho de Luz *et al.* (2013).

Os resultados do custo computacional são apresentados na Tabela 12. Luz *et al.* (2013) não relatam medidas de tempo para extração de características, apenas para treinamento e teste. No entanto, essa informação foi adicionada, visto que um sistema de auxílio à decisão deve ser composto por todas essas etapas. A melhor combinação, HOS com Bayes, que obteve uma acurácia superior aos trabalhos comparados, teve um tempo total para classificação de um batimento cardíaco de 1,74E-4s.

Uma vez que os batimentos cardíacos normais representam 88,91% de todas as amostras, considerou-se como linha de base uma acurácia de 90%. Isso significa que métodos com acurácia abaixo de 90% não são eficazes. As Figuras 27 (a) e (b) comparam a melhor combinação da proposta clássica com os resultados de Luz *et al.* (2013) na divisões de 5 e 3 classes, respectivamente. A combinação do HOS com o Bayes obtiveram acurácias de 94% e

Tabela 11 – Resultados das combinações de extratores e classificadores na base de dados MIT-BIH considerando a divisão de 3 classes com a abordagem clássica.

Classificador	Ac	N SelEs MH			VEB SelEs MH			SVEB SelEs MH		
<b>Fourier</b>										
Bayes	40,6	37,8	67,6	48,5	78,7	41,4	54,3	55,5	92,9	69,5
MLP	76,7	80,4	50,2	61,8	8,1	95,1	15,0	65,5	83,3	73,4
OPF	83,8	92,5	30,9	46,3	5,3	96,0	10,0	40,1	96,1	56,6
SVM	91,8	99,5	29,2	45,2	0,0	100,0	0,0	44,0	99,5	61,0
<b>HOS</b>										
Bayes	94,0	99,6	78,9	88,0	90,2	99,8	94,8	83,6	99,8	91,0
MLP	93,0	99,8	38,5	55,6	0,2	99,9	0,4	58,0	99,9	73,3
OPF	88,1	93,4	45,5	61,2	2,1	96,0	4,1	59,9	97,2	74,1
SVM	89,1	99,9	1,3	2,6	0,0	100,0	0,0	2,0	99,9	3,9
<b>Goertzel</b>										
Bayes	39,3	36,5	68,1	47,5	77,4	40,0	52,7	54,0	92,5	68,2
MLP	76,4	80,3	47,1	59,3	3,8	96,6	7,3	65,5	82,4	73,0
OPF	81,3	92,6	27,5	42,4	8,7	96,1	16,0	34,4	96,3	50,7
SVM	90,7	99,8	16,6	28,5	0,0	100,0	0,0	25,1	99,8	40,1
<b>SCM</b>										
Bayes	82,6	86,1	54,8	67,0	2,5	98,4	4,9	80,9	87,3	84,0
MLP	91,1	99,3	24,8	39,6	0,0	99,9	0,0	37,3	99,4	54,3
OPF	84,4	91,6	28,4	43,3	4,3	95,9	8,2	32,2	95,0	48,1
SVM	91,1	99,9	19,6	32,8	0,0	100,0	0,0	29,6	99,9	45,7
<b>SCM-Fourier</b>										
Bayes	88,6	97,1	19,4	32,3	0,5	99,1	1,0	28,7	97,9	44,4
MLP	91,7	98,9	33,4	49,9	0,7	99,5	1,3	49,8	99,3	66,4
OPF	83,5	92,0	29,0	44,0	2,7	96,0	5,3	52,7	95,6	67,9
SVM	89,7	100,0	6,5	12,3	0,0	100,0	0,0	9,8	100,0	17,9
<b>Extratores de características clássicos adotados em (LUZ et al., 2013)</b>										
<b>Chazal et al. (2004)</b>										
SVM	91,0	97,9	36,0	52,6	0,0	100,0	0,0	51,7	97,8	67,7
<b>Song et al. (2005)</b>										
SVM	89,4	100,0	3,1	6,0	0,0	100,0	0,0	04,6	100,0	8,7
<b>Yu e Chou (2008)</b>										
MLP	92,7	97,2	58,5	73,0	5,6	99,3	10,6	81,9	97,7	89,1
<b>Güler e Übeyli (2005)</b>										
SVM	89,5	99,9	05,3	100,0	0,0	100,0	0,0	08,0	99,9	14,7
<b>Yu e Chen (2007)</b>										
SVM	89,0	100,0	100,0	0,0	0,0	100,0	0,0	0,0	100,0	0,0
<b>Ye et al. (2010)</b>										
SVM	92,1	99,8	32,9	49,5	0,0	100,0	0,0	44,7	99,4	61,7

Fonte: elaborado pelo autor.

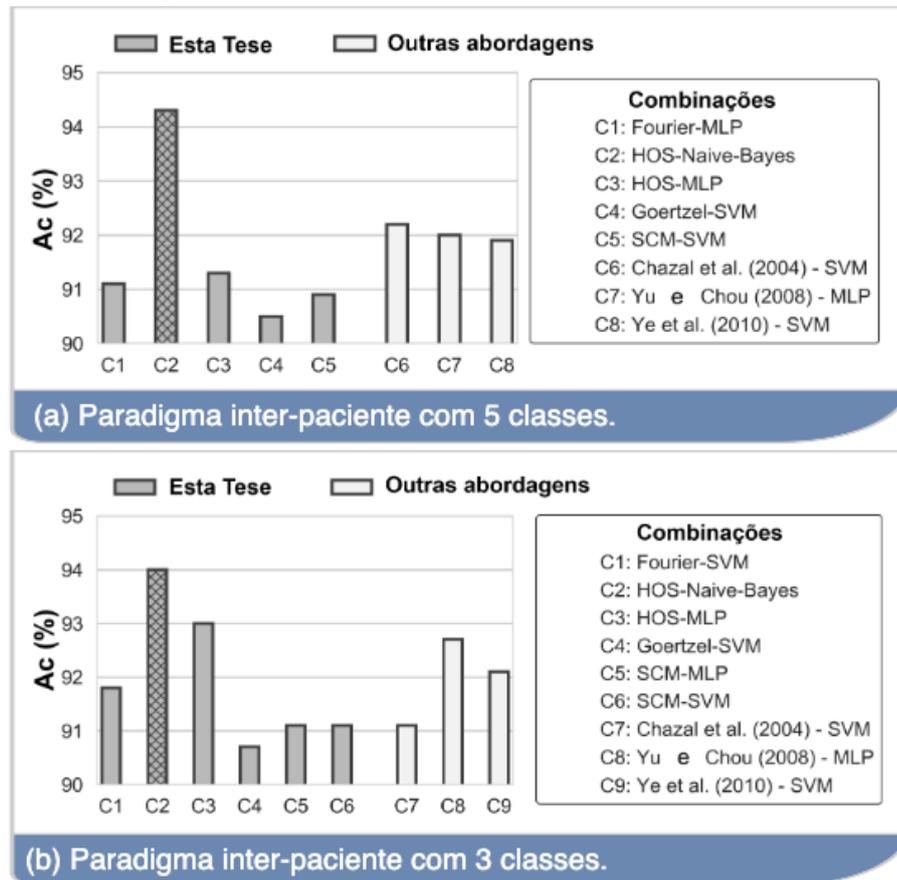
Tabela 12 – Resultados do tempo de treinamento, teste e de extração considerando a divisão de 3 classes.

<b>Classificador</b>	<b>Ac</b>	<b>Treinamento (s)</b>	<b>Teste (s)</b>	<b>Extração (s)</b>
<b>Fourier</b>				
<b>Bayes</b>	40,6	3,4E-02	9,1E-07	
<b>MLP</b>	76,7	4,7E+01	9,9E-06	
<b>OPF</b>	83,8	1,7E+02	8,3E-03	4,42E-05
<b>SVM</b>	91,8	2,7E+01	3,2E-04	
<b>HOS</b>				
<b>Bayes</b>	94,0	9,7E-03	8,8E-08	
<b>MLP</b>	93,0	8,2E+01	9,3E-06	
<b>OPF</b>	88,1	8,4E+01	3,7E-03	1,74E-04
<b>SVM</b>	89,1	1,2E+01	1,7E-04	
<b>Goertzel</b>				
<b>Bayes</b>	39,3	2,7E-02	7,5E-07	
<b>MLP</b>	76,4	3,3E+01	2,8E-06	
<b>OPF</b>	81,3	1,9E+02	8,2E-03	1,83E-03
<b>SVM</b>	90,7	3,0E+01	3,6E-04	
<b>SCM</b>				
<b>Bayes</b>	82,6	1,16E-02	1,67E-07	
<b>MLP</b>	91,1	1,91E+01	1,25E-06	
<b>OPF</b>	84,4	1,00E+02	4,00E-03	1,55E-03
<b>SVM</b>	91,1	5,74E+01	2,44E-04	
<b>SCM-Fourier</b>				
<b>Bayes</b>	88,6	1,73E-02	1,68E-07	
<b>MLP</b>	91,7	1,71E+01	1,14E-06	
<b>OPF</b>	83,5	8,27E+01	3,47E-03	1,59E-03
<b>SVM</b>	89,7	5,62E+01	2,46E-04	
<b>Extratores de características clássicos adotados em (LUZ <i>et al.</i>, 2013)</b>				
<b>Chazal et al. (2004)</b>				
<b>SVM</b>	91,0	1,9E+02	1,8E+02	-
<b>Song et al. (2005)</b>				
<b>SVM</b>	89,4	9,5E+01	5,8E+01	-
<b>Yu e Chou (2008)</b>				
<b>MLP</b>	92,7	1,8E+03	1,3E-01	-
<b>Güler e Übeyli (2005)</b>				
<b>SVM</b>	89,5	7,8E+01	6,0E+01	-
<b>Yu e Chen (2007)</b>				
<b>SVM</b>	89,0	1,1E+02	7,0E+01	-
<b>Ye et al. (2010)</b>				
<b>SVM</b>	92,1	1,6E+02	1,3E+02	-

Fonte: elaborado pelo autor.

94,3%, enquanto os melhores resultados da literatura alcançaram 92,2% e 92,7%.

Figura 27 – Comparação entre os métodos da abordagem clássica proposta nesta tese e dos trabalhos de Luz *et al.* (2013). A melhor abordagem é destacada pela barra hachurada.



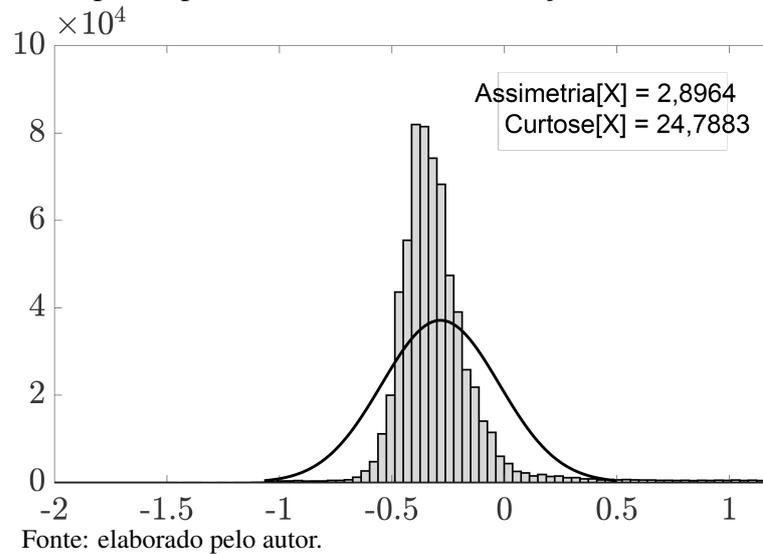
Fonte: elaborado pelo autor.

A discussão apresentada aqui pode fazer a seguinte pergunta: **Por que o extrator de características HOS tem um desempenho melhor do que os demais?.** Quando este método é usado, adota-se a hipótese de que o sinal de ECG pode ser modelado como um processo gaussiano ou, pelo menos, contém algumas dessas características. A Figura 28 exibe o histograma do registro de ECG para o paciente 101. Verifica-se que há uma semelhança entre uma distribuição normal permitindo extração de estatísticas de ordem superior.

## 4.2 Resultados da classificação de arritmia cardíaca usando Aprendizado Profundo

Os resultados com as abordagens clássicas para classificação de arritmia cardíaca usando ECG mostraram as limitações destas técnicas para resolver o problema. Diante disso, na segunda etapa da pesquisa buscou-se utilizar técnicas mais modernas e promissoras por meio do

Figura 28 – Comparação entre o histograma do registro de ECG para o paciente 101 e uma distribuição Gaussiana.



Aprendizado Profundo.

A Tabela 13 apresenta os resultados obtidos pelo método proposto nesta tese para classificação da arritmia cardíaca usando ECG. Nesse grupo de experimentos, foram adotados três paradigmas de avaliação: inter-paciente, intra-paciente e com toda a base de dados. Ademais, verificou-se a eficiência do método no problema com 5 classes, 3 classes (CHAZAL *et al.*, 2004) e com a versão binária. As métricas utilizadas são acurácia (Ac), sensibilidade (Se), especificidade (Es), valor preditivo positivo (VPP), *F1*-Score (FS) e média harmônica (MH).

Tabela 13 – Resultados do método proposto para classificação das arritmias cardíacas considerando diferentes divisões de classes e paradigmas de avaliação.

Paradigma	Ac	Se	Es	VPP	FS	MH
<b>2 classes (Normal ou Arritmia)</b>						
<b>Inter</b>	99,56±0,02	99,56±0,02	99,62±0,03	99,57±0,02	99,57±0,02	99,59±0,01
<b>Intra</b>	98,92±0,12	98,92±0,12	99,50±0,06	98,90±0,07	98,91±0,12	99,21±0,06
<b>Tudo</b>	99,16±0,01	99,16±0,01	99,57±0,02	99,16±0,01	99,16±0,01	99,37±0,01
<b>3 classes (Normal, SVEB ou VEB)</b>						
<b>Inter</b>	99,55±0,02	99,55±0,02	99,21±0,04	99,56±0,02	99,56±0,03	99,38±0,02
<b>Intra</b>	99,45±0,05	99,45±0,05	97,36±0,03	99,59±0,04	99,50±0,07	98,39±0,05
<b>Tudo</b>	99,52±0,02	99,52±0,03	99,19±0,02	99,53±0,02	99,52±0,03	99,35±0,02
<b>5 classes (Normal, SVEB, VEB, F ou Q)</b>						
<b>Inter</b>	99,45±0,03	99,45±0,03	99,20±0,02	99,46±0,04	99,46±0,03	99,32±0,03
<b>Intra</b>	98,70±0,08	98,70±0,08	82,06±0,06	98,91±0,07	98,61±0,04	89,34±0,02
<b>Tudo</b>	98,95±0,01	98,95±0,02	96,29±0,02	98,97±0,03	98,86±0,02	97,49±0,03

Fonte: elaborado pelo autor.

De acordo com os resultados exibidos na Tabela 13, observa-se que as melhores acurácias do método proposto foram 99,56%, 99,55% e 99,45% nas versões com 2, 3 e 5 classes, respectivamente. Além disso, as outras métricas calculadas a partir da matriz de confusão confirmam o bom desempenho do método no problema com classes desbalanceadas. Esses valores foram superiores à maioria das técnicas apresentadas na literatura, conforme mostra a Tabela 14.

Vale ressaltar que a grande variação das metodologias presentes nessas técnicas dificulta a comparação direta entre os resultados. Dessa forma, foram selecionadas aquelas que usam a mesma base de dados e possuem as estratégias de divisão de classe semelhantes à desta tese.

Tabela 14 – Comparação entre os melhores resultados obtidos pelo método proposto e os trabalhos relacionados usando aprendizado profundo para classificar arritmia cardíaca.

Método	# Classes	Resultados
<b>Esta Tese</b>	2	Ac: 99,56%; Se: 99,56%; Es: 99,62%; VPP: 99,57%
<b>Esta Tese</b>	3	Ac: 99,55%; Se: 99,55%; Es: 99,21%; VPP: 99,56%
<b>Esta Tese</b>	5	Ac: 99,45%; Se: 99,20%; Es: 99,46%; VPP: 99,46%
Huang <i>et al.</i> (2019)	5	Ac: 99,0%
Oh <i>et al.</i> (2019)	5	Ac: 97,32%
Xu <i>et al.</i> (2019)	5	Exp,1: Ac: 93,1% Exp, 2: Ac: 94,7% Exp, 3: Ac: 99,9% Exp, 4: Ac: 99,7%
Sannino e Pietro (2018)	2	Ac: 99,68%, Se: 99,48%, Es: 99,83%
Li <i>et al.</i> (2018)	5	VEB: Ac: 98,8%, Se: 95,5%, Es: 99,1% SVEB: Ac: 98,3%, Se: 68,7%, Es: 99,8%
Oh <i>et al.</i> (2018)	5	Ac: 98,1%, Se: 97,5%, Es: 98,7%
Yıldırım <i>et al.</i> (2018)	13 15 17	Ac: 95,20% Ac: 92,51% Ac: 91,33%
Zhai e Tin (2018)	5	VEB: Ac: 99,1%, Se: 96,4%, Es: 99,5% SVEB: Ac: 97,3%, Se: 85,3%, Es: 98,0%
Luo <i>et al.</i> (2017)	4	Patient-Esific scenario: VEB: Ac: 99,1%, Se: 93,3%, Es: 99,5% SVEB: Ac: 98,8%, Se: 71,4%, Es: 99,8% Inter-patient scenario: VEB: Ac: 95,5%, Se: 60,4%, Es: 97,9% SVEB: Ac: 96,2%, Se: 15,4%, Es: 99,3%
Kiranyaz <i>et al.</i> (2016)	5	VEB: Ac: 99%, Se: 93,9%, Es: 98,9% SVEB: Ac: 97,6%, Se: 60,3%, Es: 99,2%

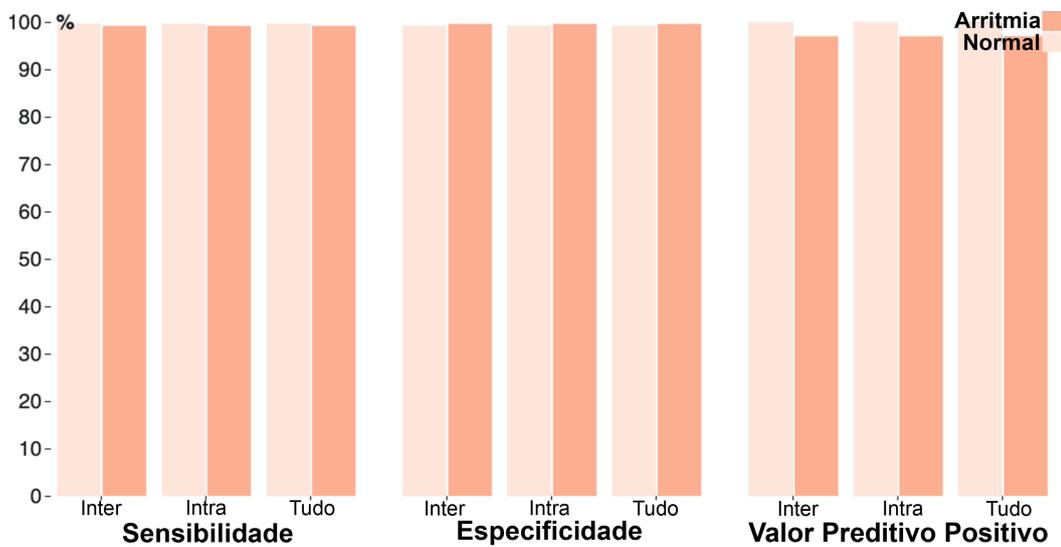
Fonte: elaborado pelo autor.

Os melhores resultados obtidos pelo método proposto são comparados, na Tabela 14, com as principais técnicas da literatura que usam aprendizado profundo para tarefa de interesse. As abordagens mais adotadas usam Rede Neural Profunda (*Deep Neural Network*, DNN), Rede Neural Convolucional (*Convolutional Neural Network*, CNN), Transformada de Fourier de Curta Duração (*Short-Time Fourier Transform*, STFT) com CNN, autoencoder convolucional e CNN com Memória Longa de Curto Prazo (*Long Short-Term Memory*, LSTM). Os pesquisadores geralmente criam novos modelos com diferentes tamanhos de camada ou tentam melhorar a

entrada no modelo (OH *et al.*, 2019). Outros trabalhos convertem o sinal em dados bidimensionais e usam modelos de processamento de imagem comuns (ZHAI; TIN, 2018; HUANG *et al.*, 2019).

As Figuras 29, 30 e 31 ilustram os resultados por classe para os diferentes paradigmas de avaliação na classificação das arritmias cardíacas. Destacou-se a sensibilidade, especificidade e valor preditivo positivo por serem as métricas mais adequadas para avaliar o desempenho do sistema de auxílio ao diagnóstico médico.

Figura 29 – Resultado do método proposto por classe para classificação binária de arritmias cardíacas no ECG usando diferentes paradigmas de avaliação.



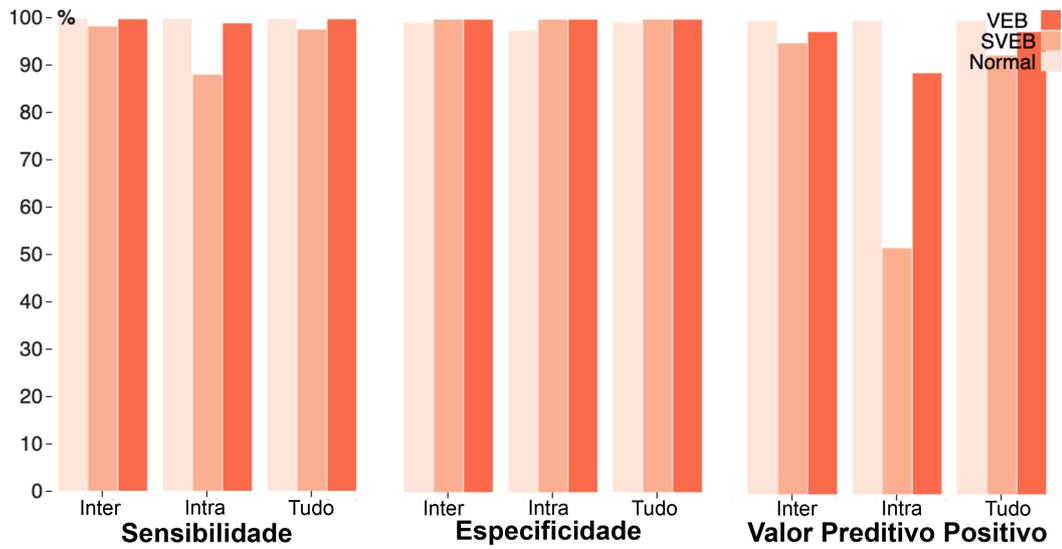
Fonte: elaborado pelo autor.

A Figura 29 confirma o excelente desempenho do método proposto na abordagem binária. Por outro lado, muitas técnicas na literatura apresentam alta acurácia, mas baixa sensibilidade e especificidade. Ou seja, os modelos possuem uma tendência a predizer a classe com maior ocorrência, o que não é relevante do ponto de vista clínico.

Verifica-se um comportamento semelhante entre as divisões com 3 e 5 classes, Figuras 30 e 31. Embora na última abordagem o método tenha apresentado dificuldade para predizer a classe Q, é importante lembrar que o seu número de amostras é consideravelmente menor que das outras classes.

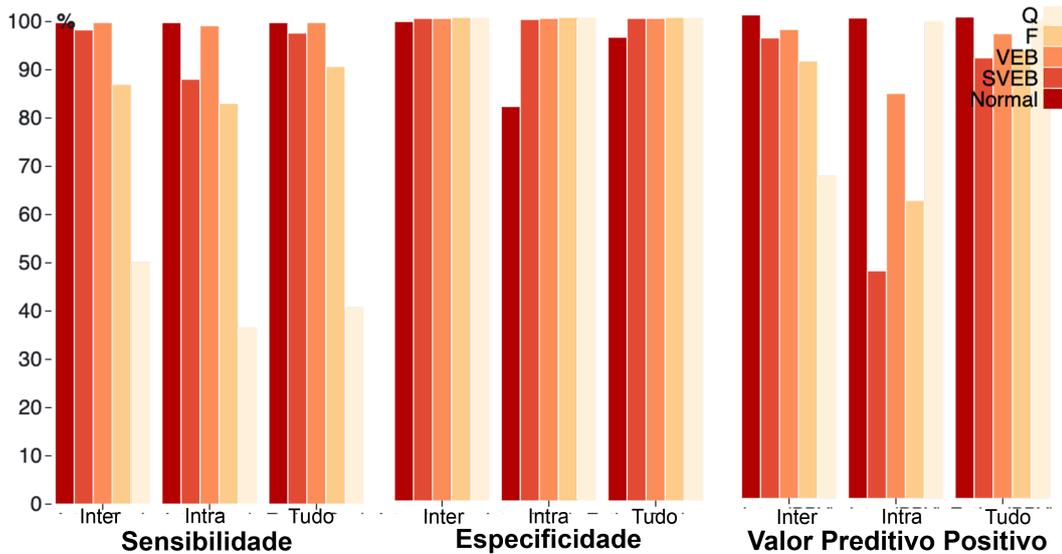
A Tabela 15 apresenta o custo computacional para treinar o método proposto e predizer um batimento cardíaco desconhecido no ECG. Além de ter um excelente desempenho no reconhecimento de anomalias cardíacas, constata-se que o método é adequado para sistemas de tempo real quando executado com Unidade de Processamento Gráfico (*Graphics Processing Unit*, GPU), apresentando uma classificação média de 5,59E-3s.

Figura 30 – Resultado do método proposto por classe na classificação de arritmias cardíacas usando a divisão em 3 classes.



Fonte: elaborado pelo autor.

Figura 31 – Resultado do método proposto por classe na classificação de arritmias cardíacas usando a divisão em 5 classes.



Fonte: elaborado pelo autor.

### 4.3 Resultados da detecção e segmentação dos batimentos no ECG

Além da classificação das arritmias cardíacas, que é o objetivo final do método proposto nesta tese, avaliou-se também a eficiência do método proposto na detecção e segmentação dos batimentos cardíacos. É importante ressaltar que um único modelo foi otimizado para realizar simultaneamente essas tarefas.

A Tabela 16 apresenta os principais trabalhos da literatura para detecção do complexo QRS usando as bases de dados MIT-BIH e QT. A maioria das abordagens usam a transformada

Tabela 15 – Custo computacional para treinar o método e prever um batimento cardíaco considerando diferentes divisões de classe e paradigmas de avaliação.

Paradigma	Tempo de Treinamento (h)	Tempo de classificação com GPU (s)	Tempo de classificação com CPU (s)	Ac (%)
<b>2 classes (Normal ou Arritmia)</b>				
<b>Inter</b>	48,76	5,65E-3	1,74E+0	99,56
<b>Intra</b>	48,50	5,58E-3	1,81E+0	98,92
<b>Tudo</b>	50,62	5,67E-3	1,75E+0	99,16
<b>3 classes (Normal, SVEB ou VEB)</b>				
<b>Inter</b>	48,21	5,55E-3	1,79E+0	99,55
<b>Intra</b>	48,98	5,44E-3	1,74E+0	99,45
<b>Tudo</b>	51,09	5,47E-3	1,88E+0	99,52
<b>5 classes (Normal, SVEB, VEB, F ou Q)</b>				
<b>Inter</b>	48,21	5,72E-3	1,80E+0	99,45
<b>Intra</b>	48,50	5,61E-3	1,79E+0	98,70
<b>Tudo</b>	50,62	5,64E-3	1,91E+0	98,95

Fonte: elaborado pelo autor.

Wavelet. Os demais trabalhos adotam a transformada de Hilbert, transformada fasorial, filtragem não linear e diferenciador de ordem fracionária digital.

De acordo com os resultados mostrados na Tabela 16, verifica-se que o método proposto obteve acurácias semelhantes às demais técnicas, 99,88%, em ambas as bases. Além disso, a sensibilidade e o valor preditivo positivo confirmam o excelente desempenho do método.

O cálculo da diferença temporal entre as marcações manuais, feitas por um especialista, e as predições previstas pelo modelo é o critério de validação na etapa de segmentação da onda no ECG. A Tabela 17 mostra a média e o desvio padrão dessas diferenças de tempo para a segmentação QRS no banco de dados QT. Os resultados mostram que o erro de segmentação obtido pelo método proposto está dentro do nível de tolerância destacado em negrito na última linha.

A Tabela 18 compara os resultados de detecção e segmentação das ondas T e P obtidos pelo método proposto nesta tese com os principais trabalhos da literatura. Os melhores resultados foram destacados em verde. Observa-se que o método apresentou excelente desempenho na detecção e segmentação em ambos os tipos de onda. Na detecção das ondas T e P, as melhores sensibilidades foram 99,85% e 99,52%, enquanto os valores preditivos positivos foram 99,92% e 99,69%. Os erros de segmentação para esses tipos de onda foram  $-0,25 \pm 13,22$  e  $-2,80 \pm 14,79$  para onda T e  $0,26 \pm 9,98$  e  $0,17 \pm 9,68$  para onda P.

Além dos indicadores tradicionais de segmentação, foi proposta e avaliada uma

Tabela 16 – Resultados da detecção do complexo QRS sobre os bancos de dados MIT-BIH e QT. *An* indica a quantidade de anotações médicas usadas na avaliação.

Método	Ac (%)	Se (%)	VPP (%)	An	VP	FP	FN
<b>MIT-BIH</b>							
<b>Esta Tese (inter)</b>	99,88	99,90	99,92	53037	52973	42	52
<b>Esta Tese (intra)</b>	99,85	99,86	99,87	75715	75599	98	106
<b>Esta Tese (tudo)</b>	99,84	99,87	99,89	109349	109175	125	145
Bashar <i>et al.</i> (2019)	99,89	99,94	99,95	109457	109391	55	66
Hossain <i>et al.</i> (2019)	99,85	99,96	99,89	109441	109395	122	46
Nayak <i>et al.</i> (2019)	99,76	99,89	99,88	109494	109442	70	52
Mourad e Fethi (2016)	99,71	99,76	99,95	106310	106055	48	259
Madeiro <i>et al.</i> (2012)	98,31	99,15	99,18	109495	108568	856	928
Martinez <i>et al.</i> (2010)	99,68	99,71	99,97	109428	109111	35	317
Ghaffari <i>et al.</i> (2009)	99,79	99,91	99,88	109428	109327	129	101
Martinez <i>et al.</i> (2004)	99,66	99,80	99,86	109428	109208	153	220
LI <i>et al.</i> (1995)	99,83	99,89	99,94	104182	104070	65	112
Hamilton e Tompkins (1986)	99,46	99,69	99,77	109267	108927	248	340
<b>Base de dados QT</b>							
<b>Esta Tese (intra)</b>	99,88	99,88	99,94	60896	60825	38	76
<b>Esta Tese (tudo)</b>	99,85	99,87	99,92	86994	86863	71	117
Hossain <i>et al.</i> (2019)	99,89	99,97	99,93	86989	86959	65	30
Qin <i>et al.</i> (2017)	99,63	99,83	99,90	86995	86848	91	147
Madeiro <i>et al.</i> (2012)	99,38	99,75	99,65	86995	86798	250	197
Martinez <i>et al.</i> (2010)	99,89	99,95	99,93	86892	86852	55	40
Ghaffari <i>et al.</i> (2009)	99,88	99,96	99,92	86892	86854	70	38
Martinez <i>et al.</i> (2004)	99,80	99,92	99,88	86892	86824	107	68
Moody e Mark (1982)	96,67	97,20	99,46	86892	84458	459	2434

Fonte: elaborado pelo autor.

Tabela 17 – Média e desvio padrão da diferença temporal, em *ms*, para a segmentação do QRS sobre o banco de dados QT.

Método	$QRS_{on}$	$QRS_{off}$
<b>Esta Tese (intra)</b>	$-0,27 \pm 12,39$	$-2,07 \pm 10,03$
<b>Esta Tese (tudo)</b>	$-0,26 \pm 12,37$	$-2,08 \pm 10,04$
Madeiro <i>et al.</i> (2012)	$2,85 \pm 9,90$	$2,83 \pm 12,26$
Martinez <i>et al.</i> (2010)	$-0,20 \pm 7,20$	$2,50 \pm 8,90$
Martinez <i>et al.</i> (2004)	$4,60 \pm 7,70$	$0,80 \pm 8,70$
Laguna <i>et al.</i> (1994)	$-1,10 \pm 8,30$	$-7,20 \pm 14,30$
Willems <i>et al.</i> (1990)	<b>6,5</b>	<b>11,6</b>

Fonte: elaborado pelo autor.

Tabela 18 – Resultados do método proposto para detecção e segmentação das ondas T e P no banco de dados QT.

<b>Onda T</b>				
Método	Se (%)	VPP (%)	$T_{pico}$ (ms)	$T_{off}$ (ms)
<b>Esta Tese (intra)</b>	99,85	99,92	$-0,25 \pm 13,22$	$-2,80 \pm 14,79$
<b>Esta Tese (tudo)</b>	99,84	99,69	$-0,29 \pm 13,21$	$-2,89 \pm 14,55$
Madeiro <i>et al.</i> (2017)	99,86	99,97	$-5,12 \pm 11,46$	$1,32 \pm 14,86$
Madeiro <i>et al.</i> (2012)	99,32	99,47	$1,4 \pm 9,0$	$2,8 \pm 15,3$
Martinez <i>et al.</i> (2010)	99,20	99,01	$5,3 \pm 12,9$	$5,8 \pm 22,7$
Ghaffari <i>et al.</i> (2009)	99,87	99,80	$0,3 \pm 4,1$	$0,8 \pm 10,7$
Martinez <i>et al.</i> (2004)	99,77	97,79	$0,2 \pm 13,9$	$-1,6 \pm 18,1$

<b>Onda P</b>				
Método	Se (%)	VPP (%)	$P_{on}$ (ms)	$P_{off}$ (ms)
<b>Esta Tese (intra)</b>	99,52	99,69	$1,61 \pm 11,17$	$1,65 \pm 10,94$
<b>Esta Tese (tudo)</b>	99,48	99,65	$0,26 \pm 9,98$	$0,17 \pm 9,68$
Madeiro <i>et al.</i> (2017)	97,76	99,73	$-4,23 \pm 14,84$	$2,26 \pm 13,14$
Martinez <i>et al.</i> (2010)	98,65	97,52	$2,6 \pm 14,5$	$0,7 \pm 14,7$
Ghaffari <i>et al.</i> (2009)	99,46	98,83	$-1,2 \pm 6,3$	$0,7 \pm 6,8$
Martinez <i>et al.</i> (2004)	99,87	91,03	$2,0 \pm 14,8$	$1,9 \pm 12,8$

Fonte: elaborado pelo autor.

métrica para verificar a diferença entre os limites das marcações dos especialistas e a onda predita na matriz  $\mathbf{M}$ , como mostra a Tabela 19. Essa diferença foi normalizada pelo tamanho do sinal predito, que é o mesmo da matriz. Observa-se que a média e o desvio padrão desses erros foram pequenos em todas as ondas. Os menores erros obtidos em módulo foram  $2,0E-4 \pm 9,7E-3$  ( $\mathbf{M}_{QRS_{on}}$ ) e  $1,3E-4 \pm 7,6E-3$  ( $\mathbf{M}_{P_{off}}$ ).

Por fim, a Figura 32 sintetiza um exemplo da saída do método proposto nesta tese em intervalos de dois diferentes registros da base de dados QT. As predições são destacadas pelos retângulos tracejados. Verifica-se que o método consegue efetuar com êxito as tarefas de detecção, segmentação e classificação de diferentes tipo de batimentos cardíacos.

#### 4.4 Considerações finais

Este capítulo apresentou os resultados dos dois grupos de experimentos propostos. No primeiro grupo analisou-se o melhor que é possível obter quanto a extração de características com uma abordagem clássica. Diante disso, foi feito um estudo do que pode ser obtido com uma abordagem de Aprendizado Profundo, gerando o segundo grupo de experimentos. O ganho

Tabela 19 – Resultado da diferença entre os limites da onda real e a predita na matriz **M** sobre a base de dados QT.

<b>Paradigma</b>	<b>M<sub>on</sub> (%)</b>	<b>M<sub>off</sub> (%)</b>
<b>Complexo QRS</b>		
Intra	-2,1E-4±9,7E-3	-1,6E-3±7,8E-3
Tudo	-2,0E-4±9,7E-3	-1,6E-3±7,8E-3
<b>Onda T</b>		
Intra	-2,0E-4±1,0E-2	-2,2E-3±1,2E-2
Tudo	-2,3E-4±1,0E-2	-2,3E-3±1,1E-2
<b>Onda P</b>		
Intra	1,3E-3±8,7E-3	2,0E-4±7,8E-3
Tudo	1,3E-3±8,5E-3	1,3E-4±7,6E-3

Fonte: elaborado pelo autor.

obtido com essa segunda estratégia é relevante, pois mesmo com complexidade maior, a acurácia média foi de 94% para 99%. Além disso, o método proposto com Aprendizado Profundo é capaz de segmentar e classificar os batimentos no ECG.

Luz *et al.* (2016) realizaram uma extensa revisão bibliográfica que organiza os estudos sobre classificação de arritmias cardíacas usando sinais ECG de acordo com a estratégia de divisão dos batimentos. Diversas pesquisas que seguem o paradigma intra-paciente obtiveram elevados valores de acurácia, mas tornam-se menos relevante do ponto de vista clínico. Isso se dá pelo fato da possibilidade dos batimentos cardíacos dos mesmos pacientes aparecem tanto no treinamento quanto no conjunto de validação.

Assim, um protocolo de avaliação se faz necessário para uma comparação mais justa entre os trabalhos. Na abordagem inter-paciente, os autores constroem seus modelos e os avaliam com batimentos cardíacos de diferentes pacientes. Chazal *et al.* (2004) propuseram um protocolo de divisão dos registros do ECG que segue esse paradigma.

Diante desta problemática, foram avaliados diferentes categorias de extratores e classificadores à luz da norma AAMI e que usavam o esquema de divisão proposto por Chazal *et al.* (2004). Os resultados foram confrontados com o trabalho de Luz *et al.* (2013) que reavaliaram alguns extratores de características que obtiveram uma acurácia próxima de 100%, mas que não adotavam nenhum protocolo de seleção de batimento cardíaco.

A semelhança entre o complexo QRS normal e uma função de densidade de proba-

Figura 32 – Resultados da saída do método proposto para intervalos de dois registros da base de dados QT.



Fonte: elaborado pelo autor.

bilidade Gaussiana, juntamente com a estratégia do HOS computar seus atributos destacaram o bom desempenho desse extrator, cuja acurácia foi de 94,3% com o classificador Bayes. Por outro lado, a combinação com maior acurácia, 92,2%, no trabalho de Luz *et al.* (2013) foi do SVM com o extrator proposto por Chazal *et al.* (2004). Além de apresentar maior acurácia, o HOS gera apenas 4 atributos, diferente dos 155 propostos no trabalho de Chazal *et al.* (2004).

A quantidade não representativa de amostras da classe Q e os valores inexpressivos alcançados pela F levaram a divisão do problema em três classes principais: N, SVEB e VEB (LLAMEDO; MARTÍNEZ, 2010). Neste cenário, verificou-se uma semelhança do HOS

em relação ao problema com cinco classes que apresentou acurácia superior aos trabalhos comparados.

Os valores das métricas obtidas a partir da matriz de confusão no primeiro grupo de estudo destacaram as deficiências destas técnicas na resolução do problema. Diante deste desafio, um novo método que usa uma abordagem de regressão com Aprendizado Profundo para sinais de ECG foi desenvolvido. O método vai além da classificação de arritmias cardíacas e através de um único modelo é capaz de segmentar os pontos fiduciais concomitantemente.

Quanto a análise da eficiência do método proposto para classificação de arritmias cardíacas, foram adotados os paradigmas inter-paciente, intra-paciente, assim como todos os batimentos da MIT-BIH. Além disso, o problema foi estudado com 5 classes, 3 classes e usando a abordagem binária. Vale ressaltar que os principais trabalhos da literatura não realizaram um estudo abrangente como o proposto nesta tese.

É importante salientar que a classe N é dominante, representando 89,46% do conjunto de dados de todos os batimentos cardíacos. Portanto, é importante confrontar a eficácia dessa classe com as demais, uma vez que um desempenho ótimo apenas na N significa uma precisão final de quase 90%. A sensibilidade e especificidade são uma das métricas que podem capturar o desempenho por cada classe.

As maiores acurácias obtidas pelo método proposto foram alcançadas com o paradigma inter-paciente nos problemas com 2, 3 e 5 classes: 99,56%, 99,55% e 99,45%. De forma semelhante, as métricas computadas a partir da matriz de confusão também foram acima de 99%, evidenciado o bom desempenho mesmo com problema desbalanceado. Por outro lado, os métodos da abordagem clássica avaliados no primeiro grupo de experimentos apresentaram tanto a acurácia geral abaixo das citadas anteriormente, como a sensibilidade e especificidade inferiores entre as classes.

O método proposto que usa uma abordagem de regressão com Aprendizado Profundo para sinais de ECG também foi avaliado na detecção e segmentação dos batimentos cardíacos. Diferente da maioria das técnicas na literatura que foram desenvolvidas para resolver cada etapa de maneira individual, o modelo proposto usa uma única saída para prever os tipos de batimento, assim como sua detecção e segmentação no sinal de ECG. A maioria das abordagens existentes usam as transformadas Wavelet, de Hilbert, fasorial, além de filtragem não linear.

Os resultados mostraram alguns desafios do método proposto. Uma vez que o modelo aprende os tipos de onda a partir das amostras do sinal ECG, ele se esforça para generalizar

os padrões de batimentos em proporções diversas. Outro ponto está relacionado a função de perda. Os erros dos batimentos pequenos são tratados do mesmo modo que nas janelas dos batimentos maiores. Assim, um pequeno erro em uma janela grande costuma ser benigno, porém um pequeno erro em uma janela pequena provoca um prejuízo maior na métrica que mede o nível de confiança. Portanto, os principais erros são os posicionamentos temporais incorretos das janelas durante a segmentação das ondas.

## 5 CONCLUSÕES E TRABALHOS FUTUROS

Nesta tese é proposto um novo método que através de uma abordagem de regressão usando Aprendizado Profundo é capaz de concomitantemente predizer os tipos de doenças cardíacas, segmentar e detectar os batimentos a partir de um sinal ECG. Inicialmente, avaliou-se diferentes técnicas clássicas de extração de características e de classificação na tarefa de interesse. Posteriormente, verificou-se a eficiência do método proposto em suas diferentes etapas. Um único modelo foi otimizado para realizar simultaneamente as etapas descritas anteriormente.

O estudo da abordagem clássica teve como objetivo buscar a combinação mais adequada de extrator de características e classificador para classificação de arritmia cardíaca usando sinais ECG. Diferentes métricas foram adotadas para selecionar aqueles que alcançaram resultados superiores ou equivalentes aos trabalhos do estado da arte, bem como o custo computacional para obtenção desses valores.

A análise do primeiro grupo de estudos mostrou que as pesquisas que não adotam o paradigma inter-paciente prejudicam a comparação entre os trabalhos da literatura, pois os mesmos não padronizaram um protocolo para as validações. Neste paradigma, intra-paciente, os batimentos cardíacos dos mesmos pacientes podem estar presentes tanto no conjunto de treinamento quanto de teste. Por outro lado, no paradigma inter-paciente os autores separaram os batimentos cardíacos de diferentes pacientes no treinamento da etapa de avaliação dos métodos.

Os extratores de características que apresentavam as maiores acurácias na literatura, mas não seguiam nenhum padrão de seleção de batimentos cardíacos foram reavaliados sobre o protocolo adotado no primeiro grupo de estudos desta tese. Dentre estes métodos, a melhor combinação foi o SVM com extrator proposto por Chazal *et al.* (2004) que conseguiu uma acurácia de 92,2%.

Por outro lado, a melhor combinação da abordagem clássica proposta nesta tese foi do extrator HOS com o classificador Bayes que obteve 94,3% de acurácia. Verifica-se também um ganho para classificação de contração ventricular prematura, classe V, cuja abordagem proposta aqui obteve 87,8% de sensibilidade, enquanto o de Luz *et al.* (2013) foi 48%. A SCM-Fourier com Bayes obteve acurácia de 92,4% e especificidade acima de 98% na maioria das classes. Além disso, a escolha do filtro de Fourier com o sinal reconstruído com 1% da componente DC na SCM teve um ganho de 2,27% em relação à sua versão com da média.

A similaridade entre o complexo QRS normal e uma função de densidade de probabilidade Gaussiana, em conjunto com o modo do HOS computar as dissimilaridades entre um

signal de entrada e uma referência Gaussiana pura favoreceram o bom desempenho deste extrator. Em contra partida, o extrator proposto por Chazal *et al.* (2004) utilizam atributos baseados em referências de tempo para medir a morfologia do batimento cardíaco no ECG. Ou seja, o método se ajusta ao conjunto de dados MIT-BIH, cujas configurações são únicas, como a frequência de amostragem. Uma vantagem também a ser destacada é o fato do HOS ter apenas 4 atributos, ao contrário dos 155 propostos pelo extrator de Chazal *et al.* (2004).

Ao investigar os resultados do problema com 3 classes no primeiro grupo de experimentos, observou-se que o HOS teve desempenho semelhante a da divisão com 5 classes. Esse extrator em conjunto com o classificador Bayes conseguiram acurácia de 94%, sensibilidade acima de 99% para os batimentos cardíacos normais e especificidade de 99% nas classes VEB e SVEB. Estes resultados foram superiores aos dos extratores com maior taxa de acerto na literatura que foram reimplementados sobre o protocolo adotado nesta tese.

O custo computacional dos métodos também foram avaliados na análise das abordagens clássicas. A melhor combinação, HOS com Bayes, que conseguiu um aumento percentual na acurácia de 2.28% para o problema de 5 classes, em relação aos trabalhos comparados, teve um tempo total para classificação de um batimento cardíaco de  $1,74E-4s$ .

Os resultados do primeiro grupo de experimentos utilizando as abordagens clássicas para classificação de arritmia cardíaca evidenciaram as deficiências destas técnicas para resolver o problema. Diante desta problemática, um novo método usando Aprendizado Profundo foi proposto. A comparação deste método com outras abordagens semelhantes na literatura foi realizada no segundo conjunto de experimentos. O método proposto foi avaliado com os paradigmas inter-paciente e intra-paciente, além de toda os batimentos da MIT-BIH. Ademais, sua eficiência foi investigada no problema com 5 classes, 3 classes e na versão binária.

O método proposto nesta tese obteve acurácias de 99,56%, 99,55% e 99,45% nas versões com 2, 3 e 5 classes, respectivamente, sobre o paradigma inter-paciente. Estes resultados foram semelhantes as das outras métricas computadas a partir da matriz de confusão, com valores superiores a 99% em todas as variações. Isto evidencia seu bom desempenho mesmo com o alto desbalanceamento das classes. É importante ressaltar que embora as acurácias das técnicas da abordagem clássica fossem elevadas, as mesmas apresentavam baixa sensibilidade ou especificidade. Ademais, os batimentos cardíacos normais representam 88,91% de todas as amostras. Portanto, uma solução clinicamente adequada para este problema demanda um bom desempenho nas métricas de avaliação para problemas desbalanceados.

Assim como nas técnicas que usam abordagem clássica para classificação de arritmias cardíacas, a grande variação das metodologias presentes nas técnicas com Aprendizado Profundo dificultou a comparação direta entre os resultados. Assim, foram escolhidos aqueles que adotaram a mesma base de dados e possuem as estratégias de divisão de classe semelhantes à desta tese. Os resultados desta comparação mostraram que o método proposto foi superior a diversos métodos da literatura. Verifica-se também que, no geral, os métodos que usam aprendizado profundo obtiveram valores melhores que os apresentados no primeiro grupo de experimentos.

Além das métricas adotadas no estudo dos problemas desbalanceados, verificou-se também o custo computacional do método proposto para prever um batimento cardíaco desconhecido no ECG. O tempo de classificação média usando GPU foi de  $5,59E-3s$ , corroborando que o mesmo é adequado para sistemas de tempo real.

Uma vez que o método proposto nesta tese consegue realizar simultaneamente a classificação das anomalias cardíacas, detectar e segmentar os batimentos, investigou-se também a eficiência do mesmo nestas etapas supracitadas. É importante destacar que um único modelo foi otimizado para realização destas etapas que ocorrem de maneira atômica.

Quanto a detecção do QRS, o método proposto obteve 99,88% de acurácia nas bases de dados MIT-BIH e QT. Esses valores foram semelhantes aos principais trabalhos da literatura que conseguiram 99,89% em ambas as bases de dados. Em relação a segmentação do QRS, o erro médio da diferença temporal do início do QRS,  $QRS_{on}$ , foi 0,26ms e o  $QRS_{off}$  foi 2,08ms. Esses valores ficaram próximos aos trabalhos do estado da arte, além de estar dentro do nível de tolerância que é de 6,5ms e 11,6ms, respectivamente. Na detecção das ondas T e P, as melhores sensibilidades foram 99,85% e 99,52%, enquanto os valores preditivos positivos foram 99,92% e 99,69%. Os erros de segmentação para esses tipos de onda foram  $-0,25 \pm 13,22$  e  $-2,80 \pm 14,79$  para onda T e  $0,26 \pm 9,98$  e  $0,17 \pm 9,68$  para onda P. O método proposto apresentou excelente desempenho na detecção e segmentação, também, nestes tipos de ondas.

Por fim, uma métrica para verificar a diferença entre os limites das marcações dos especialistas e a onda predita na matriz  $\mathbf{M}$  foi proposta para comparações futuras. Constatou-se que a média e o desvio padrão dos erros foram baixos em todas as ondas. Os menores erros obtidos em módulo foram  $2,0E-4 \pm 9,7E-3$  ( $\mathbf{M}_{QRS_{on}}$ ) e  $1,3E-4 \pm 7,6E-3$  ( $\mathbf{M}_{P_{off}}$ ).

Deste modo, resumidamente, as principais contribuições desta tese são: um novo método para classificação, detecção e segmentação de batimentos cardíacos em sinais de eletrocardiograma, que se mostrou promissor não somente na tarefa de interesse, mas também para

outros problemas semelhantes de processamento digital de sinais. Além disso, um extenso estudo foi realizado investigando a eficiência de métodos e metodologias desde abordagens clássicas até o uso de técnicas mais modernas que usam Aprendizado Profundo para classificação de arritmias cardíacas. Por fim, diversas publicações são produzidas, indicando que a comunidade acadêmica reconhece tais contribuições.

O desenvolvimento do estudo realizado nesta tese, bem como os resultados obtidos não se limitam a pesquisa com o método proposto. Neste sentido, pode-se propor os seguintes trabalhos futuros:

- avaliar a influência de outros canais de derivação, assim como a combinação dos mesmos;
- explorar o desempenho do método proposto em outras bases de dados de ECG;
- usar LSTM com valores  $K$  adaptativo no método proposto;
- avaliar outras arquiteturas como aprendizado residual e *MobileNets*;
- combinar diferentes extratores com desvios padrões semelhantes;
- calcular mais de uma matriz de distância e combinar no processo de extração de características;
- combinar extratores *handcrafted* com DL;
- empregar arquiteturas de CNN unidimensionais;
- avaliar outros mapeamentos do sinal, tais como truncamento e *reshape*;
- investigar o método em outros sinais fisiológicos, tal como o eletroencefalograma.

## REFERÊNCIAS

- ALICKOVIC, E.; SUBASI, A. Effect of multiscale PCA de-noising in ecg beat classification for diagnosis of cardiovascular diseases. **Circuits, Systems, and Signal Processing**, v. 34, n. 2, p. 513–533, FEB 2015. ISSN 0278-081X.
- ANDREÃO, R. V.; BOUDY, J. Combining wavelet transform and hidden markov models for ecg segmentation. **EURASIP Journal on Advances in Signal Processing**, Springer, v. 2007, p. 1–8, 2006.
- BASHAR, S. K.; NOH, Y.; WALKEY, A. J.; MCMANUS, D. D.; CHON, K. H. Verb: Vfcdm-based electrocardiogram reconstruction and beat detection algorithm. **IEEE Access**, IEEE, v. 7, p. 13856–13866, 2019.
- BAZI, Y.; ALAJLAN, N.; ALHICHRI, H.; MALEK, S. Domain adaptation methods for ecg classification. In: IEEE. **2013 international conference on computer medical applications (ICCMA)**. Amman, Jordan, 2013. p. 1–4.
- BERAZA, I.; ROMERO, I. Comparative study of algorithms for ecg segmentation. **Biomedical Signal Processing and Control**, v. 34, p. 166 – 173, 2017. ISSN 1746-8094.
- BERGSTRA, J.; BENGIO, Y. Random search for hyper-parameter optimization. **J. Mach. Learn. Res.**, JMLR.org, v. 13, p. 281–305, fev. 2012. ISSN 1532-4435.
- BISHOP, C. M. **Pattern recognition and machine learning**. Secaucus, NJ, USA: springer, 2006. ISBN 0387310738.
- BORTOLAN, G.; CHRISTOV, I.; PEDRYCZ, W. Hyperbox classifiers for ecg beat analysis. In: IEEE. **2007 Computers in Cardiology**. Durham, North Carolina, 2007. p. 145–148.
- BRAGA, A. de P.; FERREIRA, A. C. P. de L.; LUDERMIR, T. B. **Redes neurais artificiais: teoria e aplicações**. Rio de Janeiro, RJ, Brasil: LTC Editora, 2007.
- BRIGHAM, E. O.; BRIGHAM, E. O. **The fast Fourier transform and its applications**. Englewood Cliffs, New Jersey: prentice Hall Englewood Cliffs, NJ, 1988. v. 448.
- BUCKBERG, G. D.; NANDA, N. C.; NGUYEN, C.; KOCICA, M. J. What is the heart? anatomy, function, pathophysiology, and misconceptions. **Journal of cardiovascular development and disease**, Multidisciplinary Digital Publishing Institute, v. 5, n. 2, p. 33, 2018.
- CAPDEVIELLE, V.; SERVIERE, C.; LACOUME, J.-L. Blind separation of wide-band sources: Application to rotating machine signals. In: IEEE. **European Signal Processing Conference, 1996. EUSIPCO 1996. 8th**. Trieste, Italy, 1996. p. 1–4.
- CHAZAL, P. D.; O'DWYER, M.; REILLY, R. B. Automatic classification of heartbeats using ecg morphology and heartbeat interval features. **IEEE transactions on biomedical engineering**, IEEE, v. 51, n. 7, p. 1196–1206, 2004.
- CHEN, W.; CLARKSON, P. M.; FAN, Q. A robust sequential detection algorithm for cardiac arrhythmia classification. **IEEE Transactions on Biomedical Engineering**, IEEE, v. 43, n. 11, p. 1120–1124, 1996.
- CLIFFORD, G. D.; AZUAJE, F.; MCSHARRY, P. **Advanced methods and tools for ECG data analysis**. Norwood, Massachusetts: Artech house Boston, 2006.

- CORTES, C.; VAPNIK, V. Support-vector networks. **Machine learning**, Springer, v. 20, n. 3, p. 273–297, 1995.
- CYBENKO, G. Approximation by superpositions of a sigmoidal function. **Mathematics of control, signals and systems**, Springer, v. 2, n. 4, p. 303–314, 1989.
- DOKUR, Z.; ÖLMEZ, T. Ecg beat classification by a novel hybrid neural network. **Computer methods and programs in biomedicine**, Elsevier, v. 66, n. 2-3, p. 167–181, 2001.
- DUDA, R. O.; HART, P. E.; STORK, D. G. **Pattern classification**. New Jersey, EUA: John Wiley & Sons, 2012.
- DUMONT, J.; HERNANDEZ, A. I.; CARRAULT, G. Improving ecg beats delineation with an evolutionary optimization process. **IEEE Transactions on Biomedical Engineering**, IEEE, v. 57, n. 3, p. 607–615, 2008.
- DWYER, R. Detection of non-gaussian signals by frequency domain kurtosis estimation. In: IEEE. **Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'83**. Boston, Massachusetts, 1983. v. 8, p. 607–610.
- EITHOVEN, W. Die galvanometrische registering des menschlichen electrocardiograms, zugleich eine beurtheilung der anwedung des kapillar-electrometers in physiologie. **Journal of the American College of Cardiology**, v. 99, p. 472–80, 1903.
- ENGIN, M. Ecg beat classification using neuro-fuzzy network. **Pattern Recognition Letters**, Elsevier, v. 25, n. 15, p. 1715–1722, 2004.
- ERSOY, O. K. A comparative review of real and complex fourier-related transforms. **Proceedings of the IEEE**, IEEE, v. 82, n. 3, p. 429–447, 1994.
- FILHO, P. P. R.; PEIXOTO, S. A.; NÓBREGA, R. V. M. da; HEMANTH, D. J.; MEDEIROS, A. G.; SANGAIAH, A. K.; ALBUQUERQUE, V. H. C. de. Automatic histologically-closer classification of skin lesions. **Computerized Medical Imaging and Graphics**, Elsevier, 2018.
- FISHER, M. R. On the mathematical foundations of theoretical statistics. **Breakthroughs in Statistics**, The Royal Society, v. 222, n. 594-604, p. 309–368, 1922.
- GHAFFARI, A.; HOMAEINEZHAD, M.; AKRAMINIA, M.; ATAROD, M.; DAEVAEIHA, M. A robust wavelet-based multi-lead electrocardiogram delineation algorithm. **Medical Engineering & Physics**, v. 31, n. 10, p. 1219 – 1227, 2009. ISSN 1350-4533.
- GOERTZEL, G. An algorithm for the evaluation of finite trigonometric series. **American Math. Monthly**, v. 65, p. 34–35, 1958.
- GOLDBERGER, A. **Goldberger's Clinical Electrocardiography**. United States: Elsevier, 2018.
- GOLDBERGER, A. L.; AMARAL, L. A.; GLASS, L.; HAUSDORFF, J. M.; IVANOV, P. C.; MARK, R. G.; MIETUS, J. E.; MOODY, G. B.; PENG, C.-K.; STANLEY, H. E. Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. **Circulation**, Am Heart Assoc, v. 101, n. 23, p. e215–e220, 2000.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. Cambridge, Massachusetts: MIT Press, 2016.

GÜLER, İ.; ÜBEYLI, E. D. Ecg beat classifier designed by combined neural network model. **Pattern recognition**, Elsevier, v. 38, n. 2, p. 199–208, 2005.

HAMILTON, P. S.; TOMPKINS, W. J. Quantitative investigation of qrs detection rules using the mit/bih arrhythmia database. **IEEE transactions on biomedical engineering**, IEEE, n. 12, p. 1157–1165, 1986.

HAYKIN, S. Neural networks: principles and practice. **Bookman**, 2001.

HORNIK, K. Approximation capabilities of multilayer feedforward networks. **Neural Networks**, v. 4, n. 2, p. 251 – 257, 1991. ISSN 0893-6080.

HOSSAIN, M. B.; BASHAR, S. K.; WALKEY, A. J.; MCMANUS, D. D.; CHON, K. H. An accurate qrs complex and p wave detection in ecg signals using complete ensemble empirical mode decomposition with adaptive noise approach. **IEEE ACCESS**, v. 7, p. 128869–128880, 2019. ISSN 2169-3536.

HUANG, H.; LIU, J.; ZHU, Q.; WANG, R.; HU, G. A new hierarchical method for inter-patient heartbeat classification using random projections and rr interval. **Biomedical engineering online**, Springer, v. 13, n. 1, p. 90, 2014.

HUANG, J.; CHEN, B.; YAO, B.; HE, W. Ecg arrhythmia classification using stft-based spectrogram and convolutional neural network. **IEEE ACCESS**, v. 7, p. 92871–92880, 2019. ISSN 2169-3536.

KALLAS, M.; FRANCIS, C.; KANAAN, L.; MERHEB, D.; HONEINE, P.; AMOUD, H. Multi-class svm classification combined with kernel pca feature extraction of ecg signals. In: IEEE. **2012 19th International Conference on Telecommunications (ICT)**. Jounieh, Lebanon, 2012. p. 1–5.

KANG, W.; BYUN, K.; KANG, H.-G. Detection of fiducial points in ecg waves using iteration based adaptive thresholds. In: IEEE. **2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)**. Milano, Italy, 2015. p. 2721–2724.

KARIMIFARD, S.; AHMADIAN, A.; KHOSHNEVISAN, M.; NAMBAKHSI, M. Morphological heart arrhythmia detection using hermitian basis functions and knn classifier. In: IEEE. **2006 International Conference of the IEEE Engineering in Medicine and Biology Society**. New York, NY, 2006. p. 1367–1370.

KIM, J.; SHIN, H. S.; SHIN, K.; LEE, M. Robust algorithm for arrhythmia classification in ecg using extreme learning machine. **Biomedical engineering online**, Springer, v. 8, n. 1, p. 31, 2009.

KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. In: BENGIO, Y.; LECUN, Y. (Ed.). **3rd International Conference on Learning Representations**. San Diego, USA: ICLR, 2015.

KIRANYAZ, S.; INCE, T.; GABBOUJ, M. Real-time patient-specific ecg classification by 1-d convolutional neural networks. **IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING**, v. 63, n. 3, p. 664–675, MAR 2016. ISSN 0018-9294.

KOHLER, B.-U.; HENNIG, C.; ORGLMEISTER, R. The principles of software qrs detection. **IEEE Engineering in Medicine and biology Magazine**, IEEE, v. 21, n. 1, p. 42–57, 2002.

KORÜREK, M.; NIZAM, A. A new arrhythmia clustering technique based on ant colony optimization. **Journal of Biomedical Informatics**, Elsevier, v. 41, n. 6, p. 874–881, 2008.

KUTLU, Y.; KUNTALP, D. Feature extraction for ecg heartbeats using higher order statistics of wpd coefficients. **Computer methods and programs in biomedicine**, Elsevier, v. 105, n. 3, p. 257–267, 2012.

LAGERHOLM, M.; PETERSON, C.; BRACCINI, G.; EDENBRANDT, L.; SORNMO, L. Clustering ecg complexes using hermite functions and self-organizing maps. **IEEE Transactions on Biomedical Engineering**, IEEE, v. 47, n. 7, p. 838–848, 2000.

LAGUNA, P.; JANÉ, R.; CAMINAL, P. Automatic detection of wave boundaries in multilead ecg signals: Validation with the cse database. **Computers and biomedical research**, Elsevier, v. 27, n. 1, p. 45–60, 1994.

LAGUNA, P.; MARK, R. G.; GOLDBERG, A.; MOODY, G. B. A database for evaluation of algorithms for measurement of qt and other waveform intervals in the ecg. In: IEEE. **Computers in cardiology 1997**. Lund, Sweden, 1997. p. 673–676.

LANATA, A.; VALENZA, G.; MANCUSO, C.; SCILINGO, E. P. Robust multiple cardiac arrhythmia detection through bispectrum analysis. **Expert Systems with Applications**, Elsevier, v. 38, n. 6, p. 6798–6804, 2011.

LANNOY, G.; FRANÇOIS, D.; DELBEKE, J.; VERLEYSSEN, M. Weighted svms and feature relevance assessment in supervised heart beat classification. In: SPRINGER. **International Joint Conference on Biomedical Engineering Systems and Technologies**. Valletta, Malta, 2010. p. 212–223.

LANNOY, G. D.; FRANÇOIS, D.; DELBEKE, J.; VERLEYSSEN, M. Weighted conditional random fields for supervised interpatient heartbeat classification. **IEEE Transactions on Biomedical Engineering**, IEEE, v. 59, n. 1, p. 241–247, 2011.

LAPAN, M. **Deep Reinforcement Learning Hands-On: Apply modern RL methods, with deep Q-networks, value iteration, policy gradients, TRPO, AlphaGo Zero and more**. Birmingham, United Kingdom: Packt Publishing Ltd, 2018.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, Nature Publishing Group, v. 521, n. 7553, p. 436–444, 2015.

LECUN, Y.; BOTTOU, L.; BENGIO, Y.; HAFFNER, P. Gradient-based learning applied to document recognition. **PROCEEDINGS OF THE IEEE**, v. 86, n. 11, p. 2278–2324, NOV 1998. ISSN 0018-9219.

LECUN, Y.; TOURESKY, D.; HINTON, G.; SEJNOWSKI, T. A theoretical framework for back-propagation. In: CMU, PITTSBURGH, PA: MORGAN KAUFMANN. **Proceedings of the 1988 connectionist models summer school**. San Matedo, Ca, 1988. v. 1, p. 21–28.

LI, C.; ZHENG, C.; TAI, C. Detection of ecg characteristic points using wavelet transforms. **IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING**, v. 42, n. 1, p. 21–28, JAN 1995. ISSN 0018-9294.

LI, F.; MENG, G.; YE, L.; CHEN, P. Wavelet transform-based higher-order statistics for fault diagnosis in rolling element bearings. **Journal of Vibration and Control**, SAGE Publications Sage UK: London, England, v. 14, n. 11, p. 1691–1709, 2008.

LI, H.; BOULANGER, P. A Survey of Heart Anomaly Detection Using Ambulatory Electrocardiogram (ECG). **SENSORS**, MDPI, ST ALBAN-ANLAGE 66, CH-4052 BASEL, SWITZERLAND, 20, n. 5, MAR 2020.

LI, Y.; PANG, Y.; WANG, J.; LI, X. Patient-specific ecg classification by deeper cnn from generic to dedicated. **Neurocomputing**, v. 314, p. 336 – 346, 2018. ISSN 0925-2312.

LIN, C.; MAILHES, C.; TOURNERET, J.-Y. P-and t-wave delineation in ecg signals using a bayesian approach and a partially collapsed gibbs sampler. **IEEE transactions on biomedical engineering**, IEEE, v. 57, n. 12, p. 2840–2849, 2010.

LIN, C.-C.; YANG, C.-M. Heartbeat classification using normalized rr intervals and morphological features. **Mathematical Problems in Engineering**, Hindawi, v. 2014, 2014.

LLAMEDO, M.; MARTÍNEZ, J. P. Heartbeat classification using feature selection driven by database generalization criteria. **IEEE Transactions on Biomedical Engineering**, IEEE, v. 58, n. 3, p. 616–625, 2010.

LUO, K.; LI, J.; WANG, Z.; CUSCHIERI, A. Patient-specific deep architectural model for ecg classification. **Journal of healthcare engineering**, Hindawi, v. 2017, 2017.

LUZ, E. J. d. S.; SCHWARTZ, W. R.; CÁMARA-CHÁVEZ, G.; MENOTTI, D. Ecg-based heartbeat classification for arrhythmia detection: A survey. **Computer methods and programs in biomedicine**, Elsevier, v. 127, p. 144–164, 2016.

LUZ, E. J. S.; NUNES, T. M.; ALBUQUERQUE, V. H. C. de; PAPA, J. P.; MENOTTI, D. Ecg arrhythmia classification based on optimum-path forest. **Expert Systems with Applications**, v. 40, n. 9, p. 3561 – 3573, 2013. ISSN 0957-4174.

MADEIRO, J. P.; CORTEZ, P. C.; MARQUES, J. A.; SEISDEDOS, C. R.; SOBRINHO, C. R. An innovative approach of qrs segmentation based on first-derivative, hilbert and wavelet transforms. **Medical Engineering & Physics**, v. 34, n. 9, p. 1236 – 1246, 2012. ISSN 1350-4533.

MADEIRO, J. P. do V.; CORTEZ, P. C.; FILHO, J. M. D. S. M.; BRAYNER, A. R. A. **Developments and Applications for ECG Signal Processing: Modeling, Segmentation, and Pattern Recognition**. Fortaleza, Brazil: Academic Press, 2018.

MADEIRO, J. P. V.; SANTOS, E. M. B.; CORTEZ, P. C.; FELIX, J. H. S.; SCHLINDWEIN, F. S. Evaluating gaussian and rayleigh-based mathematical models for t and p-waves in ecg. **IEEE LATIN AMERICA TRANSACTIONS**, v. 15, n. 5, p. 843–853, MAY 2017. ISSN 1548-0992.

MAR, T.; ZAUNSEDER, S.; MARTÍNEZ, J. P.; LLAMEDO, M.; POLL, R. Optimization of ecg classification by means of feature selection. **IEEE transactions on Biomedical Engineering**, IEEE, v. 58, n. 8, p. 2168–2177, 2011.

MARINHO, L. B.; NASCIMENTO, N. de M.; SOUZA, J. W. M.; GURGEL, M. V.; Rebouças Filho, P. P.; de Albuquerque, V. H. C. A novel electrocardiogram feature extraction approach for cardiac arrhythmia classification. **Future Generation Computer Systems**, v. 97, p. 564 – 577, 2019. ISSN 0167-739X.

MARTIN, H.; HONARVAR, F. Application of statistical moments to bearing failure detection. **Applied acoustics**, Elsevier, v. 44, n. 1, p. 67–77, 1995.

MARTINEZ, A.; ALCARAZ, R.; RIETA, J. J. Application of the phasor transform for automatic delineation of single-lead ecg fiducial points. **PHYSIOLOGICAL MEASUREMENT**, v. 31, n. 11, p. 1467–1485, NOV 2010. ISSN 0967-3334.

MARTINEZ, J.; ALMEIDA, R.; OLMOS, S.; ROCHA, A.; LAGUNA, P. A wavelet-based ecg delineator: Evaluation on standard databases. **IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING**, v. 51, n. 4, p. 570–581, APR 2004. ISSN 0018-9294.

MENDEL, J. M. Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications. **Proceedings of the IEEE**, v. 79, n. 3, p. 278–305, March 1991. ISSN 0018-9219.

MISHRA, A. K.; RAGHAV, S. Local fractal dimension based ecg arrhythmia classification. **Biomedical Signal Processing and Control**, Elsevier, v. 5, n. 2, p. 114–123, 2010.

MOODY, G. B.; MARK, R. G. Development and evaluation of a 2-lead ecg analysis program. **Computers in cardiology**, v. 9, p. 39–44, 1982.

MOODY, G. B.; MARK, R. G. The impact of the mit-bih arrhythmia database. **IEEE Engineering in Medicine and Biology Magazine**, IEEE, v. 20, n. 3, p. 45–50, 2001.

MOURAD, K.; FETHI, B. R. Efficient automatic detection of QRS complexes in ECG signal based on reverse biorthogonal wavelet decomposition and nonlinear filtering. **Measurement: Journal of the International Measurement Confederation**, v. 94, p. 663–670, dec 2016. ISSN 02632241.

MURAT, F.; YILDIRIM, O.; TALO, M.; BALOGLU, U. B.; DEMIR, Y.; ACHARYA, U. R. Application of deep learning techniques for heartbeats detection using ecg signals-analysis and review. **Computers in Biology and Medicine**, v. 120, p. 103726, 2020. ISSN 0010-4825.

MURTY, M. N.; RAGHAVA, R. **Support Vector Machines and Perceptrons: Learning, Optimization, Classification, and Application to Social Networks**. Cham: Springer, 2016.

NAJARIAN, K.; SPLINTER, R. **Biomedical signal and image processing**. Florida, New York: Taylor & Francis, 2012.

NASCIMENTO, N. M. M.; MARINHO, L. B.; PEIXOTO, S. A.; MADEIRO, J. P. do V.; ALBUQUERQUE, V. H. C. de; FILHO, P. P. R. Heart arrhythmia classification based on statistical moments and structural co-occurrence. **Circuits, Systems, and Signal Processing**, Springer, v. 39, n. 2, p. 631–650, 2020.

NAYAK, C.; SAHA, S. K.; KAR, R.; MANDAL, D. An Efficient and Robust Digital Fractional Order Differentiator Based ECG Pre-Processor Design for QRS Detection. **IEEE Transactions on Biomedical Circuits and Systems**, v. 13, n. 4, p. 682–696, aug 2019. ISSN 19409990.

OH, S. L.; NG, E. Y.; TAN, R. S.; ACHARYA, U. R. Automated diagnosis of arrhythmia using combination of cnn and lstm techniques with variable length heart beats. **Computers in Biology and Medicine**, v. 102, p. 278 – 287, 2018. ISSN 0010-4825.

OH, S. L.; NG, E. Y.; TAN, R. S.; ACHARYA, U. R. Automated beat-wise arrhythmia diagnosis using modified u-net on extended electrocardiographic recordings with heterogeneous arrhythmia types. **Computers in Biology and Medicine**, v. 105, p. 92 – 101, 2019. ISSN 0010-4825.

OSOWSKI, S.; LINH, T. H. Ecg beat recognition using fuzzy hybrid neural network. **IEEE Transactions on Biomedical Engineering**, IEEE, v. 48, n. 11, p. 1265–1271, 2001.

ÖZBAY, Y.; TEZEL, G. A new method for classification of ecg arrhythmias using neural network with adaptive activation function. **Digital Signal Processing**, Elsevier, v. 20, n. 4, p. 1040–1049, 2010.

PAPA, J. P.; FALCAO, A. X.; SUZUKI, C. T. Supervised pattern classification based on optimum-path forest. **International Journal of Imaging Systems and Technology**, Wiley Online Library, v. 19, n. 2, p. 120–131, 2009.

PARK, K. S.; CHO, B. H.; LEE, D. H.; SONG, S. H.; LEE, J. S.; CHEE, Y. J.; KIM, I. Y.; KIM, S. I. Hierarchical support vector machine based heartbeat classification using higher order statistics and hermite basis function. In: IEEE. **2008 Computers in Cardiology**. Bologna, Itália, 2008. p. 229–232. ISBN 978-1-4244-3706-1. 35th Annual Conference on Computers in Cardiology, Bologna, ITALY, SEP 14-17, 2008.

PASTORE, C.; PINHO, J.; PINHO, C.; SAMESIMA, N.; FILHO, H. P.; KRUSE, J.; PAIXÃO, A.; PÉREZ-RIERA, A.; RIBEIRO, A.; OLIVEIRA, C. III diretrizes da sociedade brasileira de cardiologia sobre análise e emissão de laudos eletrocardiográficos. **Arquivos Brasileiros de Cardiologia**, SciELO Brasil, v. 106, n. 4, p. 1–23, 2016.

PEIXOTO, S. A.; FILHO, P. P. R. Neurologist-level classification of stroke using a structural co-occurrence matrix based on the frequency domain. **Computers & Electrical Engineering**, v. 71, p. 398 – 407, 2018. ISSN 0045-7906.

PLAWIAK, P. Novel methodology of cardiac health recognition based on ecg signals and evolutionary-neural system. **Expert Systems with Applications**, Elsevier, v. 92, p. 334–349, 2018.

QIN, Q.; LI, J.; YUE, Y.; LIU, C. An adaptive and time-efficient ecg r-peak detection algorithm. **Journal of Healthcare Engineering**, Hindawi, v. 2017, 2017.

RAMALHO, G. L. B.; FERREIRA, D. S.; FILHO, P. P. R.; MEDEIROS, F. N. S. de. Rotation-invariant feature extraction using a structural co-occurrence matrix. **Measurement**, Elsevier, v. 94, p. 406–415, 2016.

REDMON, J.; DIVVALA, S.; GIRSHICK, R.; FARHADI, A. You only look once: Unified, real-time object detection. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. Las Vegas, NV: RSIP Vision, 2016. v. 1, p. 779–788.

REESE-PETERSEN, A. L.; OLESEN, M. S.; KARSDAL, M.; SVENDSEN, J. H.; GENOVESE, F. Atrial fibrillation and cardiac fibrosis: A review on the potential of extracellular matrix proteins as biomarkers. **Matrix Biology**, Elsevier, 2020.

RUSSELL, S. J.; NORVIG, P. **Artificial intelligence: a modern approach**. Englewood Cliffs, New Jersey: Pearson Education Limited,, 2016.

SAMANTA, B.; AL-BALUSHI, K. Artificial neural network based fault diagnostics of rolling element bearings using time-domain features. **Mechanical systems and signal processing**, Elsevier, v. 17, n. 2, p. 317–328, 2003.

SANNINO, G.; PIETRO, G. D. A deep learning approach for ecg-based heartbeat classification for arrhythmia detection. **Future Generation Computer Systems-the International Journal Of Science**, v. 86, p. 446–455, SEP 2018. ISSN 0167-739X.

SILVA, S. P. P.; MARINHO, L. B.; ALMEIDA, J. S.; FILHO, P. P. R. A novel approach for mobile robot localization in topological maps using classification with reject option from structural co-occurrence matrix. In: SPRINGER. **International Conference on Computer Analysis of Images and Patterns**. Ystad, Sweden, 2017. p. 3–15.

SONG, M. H.; LEE, J.; CHO, S. P.; LEE, K. J.; YOO, S. K. Support vector machine based arrhythmia classification using reduced features. **INTERNATIONAL JOURNAL OF CONTROL AUTOMATION AND SYSTEMS**, Korean Institute of Electrical Engineers, 2005.

SORIA, M. L.; MARTÍNEZ, J. Analysis of multidomain features for ecg classification. In: IEEE. **2009 36th Annual Computers in Cardiology Conference (CinC)**. Park City, Utah, USA, 2009. p. 561–564.

SOUZA, J. W.; ALVES, S. S.; REBOUÇAS, E. d. S.; ALMEIDA, J. S.; FILHO, P. P. R. A new approach to diagnose parkinson's disease using a structural cooccurrence matrix for a similarity analysis. **Computational intelligence and neuroscience**, Hindawi, v. 2018, 2018.

SZEGEDY, C.; LIU, W.; JIA, Y.; SERMANET, P.; REED, S.; ANGUELOV, D.; ERHAN, D.; VANHOUCHE, V.; RABINOVICH, A. Going deeper with convolutions. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. Puerto Rico, USA: IEEE, 2015. p. 1–9.

TORREY, L.; SHAVLIK, J. Transfer learning. In: **Handbook of research on machine learning applications and trends: algorithms, methods, and techniques**. EUA: IGI global, 2010. p. 242–264.

TRAN, H. L.; VUONG, H. N. *et al.* Multiple neural network integration using a binary decision tree to improve the ecg signal recognition accuracy. **International Journal of Applied Mathematics and Computer Science**, Sciendo, v. 24, n. 3, p. 647–655, 2014.

VAPNIK, V. N. An overview of statistical learning theory. **IEEE transactions on neural networks**, IEEE, v. 10, n. 5, p. 988–999, 1999.

WANG, J.; CHEN, Q.; CHEN, Y. Rbf kernel based support vector machine with universal approximation and its application. In: YIN, F.-L.; WANG, J.; GUO, C. (Ed.). **Advances in Neural Networks – ISNN 2004**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004. p. 512–517. ISBN 978-3-540-28647-9.

WELLENS, H. J. **The conduction system of the heart: structure, function and clinical implications**. Berlin, Germany: Springer Science & Business Media, 2012.

WEN, C.; LIN, T.; CHANG, K.; HUANG, C. Classification of ecg complexes using self-organizing cmac. **Measurement**, Elsevier, v. 42, n. 3, p. 399–407, 2009.

WHO. **Managing epidemics: key facts about major deadly diseases**. Geneva, Switzerland: World Health Organization, 2018.

WILLEMS, J.; ARNAUD, P.; BEMMEL, J. V.; DEGANI, R.; MACFARLANE, P.; ZYWIETZ, C. *et al.* Common standards for quantitative electrocardiography: goals and main results. **Methods of information in medicine**, Schattauer GmbH, v. 29, n. 04, p. 263–271, 1990.

WILSON, F. N.; JOHNSTON, F. D.; MACLEOD, A. G.; BARKER, P. S. Electrocardiograms that represent the potential variations of a single electrode. **American Heart Journal**, Elsevier, v. 9, n. 4, p. 447–458, 1934.

XU, S. S.; MAK, M.; CHEUNG, C. Towards end-to-end ecg classification with raw signal extraction and deep neural networks. **IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS**, v. 23, n. 4, p. 1574–1584, JUL 2019. ISSN 2168-2194.

YE, C.; COIMBRA, M. T.; KUMAR, B. V. Arrhythmia detection and classification using morphological and dynamic features of ecg signals. In: IEEE. **2010 Annual International Conference of the IEEE Engineering in Medicine and Biology**. Buenos Aires, Argentina, 2010. p. 1918–1921.

YE, C.; KUMAR, B. V.; COIMBRA, M. T. Combining general multi-class and specific two-class classifiers for improved customized ecg heartbeat classification. In: IEEE. **Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)**. Tsukuba, Japan, 2012. p. 2428–2431.

YILDIRIM, Ö.; PŁAWIAK, P.; TAN, R.-S.; ACHARYA, U. R. Arrhythmia detection using deep convolutional neural network with long duration ecg signals. **Computers in biology and medicine**, Elsevier, v. 102, p. 411–420, 2018.

YU, S.; CHEN, Y. Electrocardiogram beat classification based on wavelet transformation and probabilistic neural network. **Pattern Recognition Letters**, Elsevier, v. 28, n. 10, p. 1142–1150, 2007.

YU, S.; CHOU, K. Integration of independent component analysis and neural networks for ecg beat classification. **Expert Systems with Applications**, Elsevier, v. 34, n. 4, p. 2841–2846, 2008.

ZHAI, X.; TIN, C. Automated ecg classification using dual heartbeat coupling based on convolutional neural network. **IEEE Access**, IEEE, v. 6, p. 27465–27472, 2018.

ZHANG, Z.; LUO, X. Heartbeat classification using decision level fusion. **Biomedical Engineering Letters**, Springer, v. 4, n. 4, p. 388–395, 2014.

ZHOU, S.; TAN, B. Electrocardiogram soft computing using hybrid deep learning cnn-elm. **Applied Soft Computing**, v. 86, p. 105778, 2020. ISSN 1568-4946.

ZIPES, M. D. D. **Electrocardiography of Arrhythmias: A Comprehensive Review E-Book: A Companion to Cardiac Electrophysiology**. Philadelphia, Pensilvânia: Elsevier Health Sciences, 2012. ISBN 9781455700851.