

Universidade de Brasília – UnB
Campus Gama – FGA
Programa de Pós-Graduação em Engenharia Biomédica

**ABORDAGENS PARA DETECÇÃO DE
SÍNDROME METABÓLICA POR MEIO DE
SINAIS DE ELETROCARDIOGRAFIA**

Edilaine Gonçalves Costa de Faria

Orientador: Dr. CRISTIANO JACQUES MIOSSO
Coorientador: Dr. EULER DE VILHENA GARCIA



UNB – UNIVERSIDADE DE BRASÍLIA

FGA – FACULDADE GAMA

PROGRAMA DE PÓS-GRADUAÇÃO

**ABORDAGENS PARA DETECÇÃO DE
SÍNDROME METABÓLICA POR MEIO DE
SINAIS DE ELETROCARDIOGRAFIA**

Edilaine Gonçalves Costa de Faria

ORIENTADOR: CRISTIANO JACQUES MIOSSO

CO-ORIENTADOR: EULER DE VILHENA GARCIA

**DISSERTAÇÃO DE MESTRADO EM
ENGENHARIA BIOMÉDICA**

PUBLICAÇÃO: 111A/2019

BRASÍLIA/DF, JUNHO DE 2019

**UNIVERSIDADE DE BRASÍLIA
FACULDADE DO GAMA
ENGENHARIA BIOMÉDICA**

**"ABORDAGENS PARA DETECÇÃO DE SÍNDROME METABÓLICA
POR MEIO DE SINAIS DE ELETROCARDIOGRAFIA."**

EDILAINE GONÇALVES COSTA DE FARIA

**DISSERTAÇÃO DE MESTRADO SUBMETIDA À FACULDADE UNB GAMA DA
UNIVERSIDADE DE BRASÍLIA, COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA
A OBTENÇÃO DO TÍTULO DE MESTRE EM ENGENHARIA BIOMÉDICA.**

APROVADA POR:



**Prof. Dr. Cristiano Jacques Miosso; FGA/UnB
(ORIENTADOR)**



**Prof. Dr. Euler de Vilhena Garcia; FGA / UnB
(COORIENTADOR)**



**Prof. Dra. Marcella Lemos Brettas Carneiro FGA / UnB
(EXAMINADOR INTERNO)**



**Prof. Dr. Jefferson Luiz Brum Marques; UFSC
(EXAMINADOR EXTERNO)**

Brasília, 21 de junho de 2019

FICHA CATALOGRÁFICA

FARIA, EDILAINE GONÇALVES COSTA DE

Abordagens para Detecção de Síndrome Metabólica por meio de Sinais de Eletrocardiografia [Distrito Federal], 2019.

67p., 210 × 297 mm (FGA/UnB Gama, Mestrado em Engenharia Biomédica, 2019).

Dissertação de Mestrado em Engenharia Biomédica, Faculdade UnB Gama, Programa de Pós-Graduação em Engenharia Biomédica.

- | | |
|-----------------------|----------------------------|
| 1. Eletrocardiografia | 2. Processamento de Sinais |
| 3. Classificadores | 4. Síndrome Metabólica |
| I. FGA UnB/UnB. | II. Título (série) |

REFERÊNCIA

FARIA, EDILAINE GONÇALVES COSTA DE (2019). Abordagens para Detecção de Síndrome Metabólica por meio de Sinais de Eletrocardiografia. Dissertação de mestrado em engenharia biomédica, Publicação 111A/2019, Programa de Pós-Graduação, Faculdade UnB Gama, Universidade de Brasília, Brasília, DF, 67p.

CESSÃO DE DIREITOS

AUTOR: Edilaine Gonçalves Costa de Faria

TÍTULO: Abordagens para Detecção de Síndrome Metabólica por meio de Sinais de Eletrocardiografia

GRAU: Mestre

ANO: 2019

É concedida à Universidade de Brasília permissão para reproduzir cópias desta dissertação de mestrado e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte desta dissertação de mestrado pode ser reproduzida sem a autorização por escrito do autor.

edilaine@aluno.unb.br

Brasília, DF – Brasil

AGRADECIMENTOS

A Deus por ter me dado força e me guiado durante realização do mestrado.

À Universidade de Brasília (UnB), especialmente ao Programa de Pós Graduação em Engenharia Biomédica, pela oportunidade e por todo conhecimento adquirido.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), pela bolsa de estudos concedida.

Ao meu orientador Dr^o Cristiano Jacques Miosso, pelas contribuições, apoio e conhecimentos transmitidos durante a realização do mestrado.

Ao meu coorientador Dr^o Euler de Vilhena Garcia, pelas contribuições e sugestões durante o desenvolvimento desta pesquisa.

Aos professores da banca pelas sugestões.

E, por fim, aos familiares e amigos pelo incentivo e apoio durante esse período.

RESUMO

O conjunto de fatores da Síndrome Metabólica (SM), que está associado principalmente à resistência à insulina e à presença da diabetes, provoca alterações fisiológicas que podem ocasionar prejuízos à saúde e qualidade de vida dos portadores. Estudos científicos associam essas alterações ao risco cardiovascular. Esses estudos apontam que parâmetros extraídos dos sinais de Eletrocardiografia (ECG) podem ser associados à presença da SM. Apesar disso, não foram encontrados artigos científicos relacionados à detecção de SM utilizando sinais de ECG. Logo, é relevante investigar elementos relacionados ao sistema cardiovascular que indiquem quadros suspeitos de SM, com o intuito de desenvolver estratégias de detecção da SM de forma objetiva e precoce por meio de sinais de ECG. Sendo assim, em virtude dessa lacuna científica, a proposta deste trabalho consistiu em desenvolver e avaliar ferramentas de detecção de SM utilizando sinais de ECG.

Para avaliar com que exatidão e precisão sistemas classificadores permitem detectar a SM a partir dos sinais de ECG, foram adotados os seguintes procedimentos. Inicialmente foram detectados os picos Q, R e S dos sinais de ECG para extrair características temporais dos sinais. Essas características foram utilizadas no treinamento e validação de sistemas classificadores para detecção da SM, utilizando duas técnicas, denominadas Máquinas de Vetores Suporte (SVM, do inglês *Support Vector Machine*) e *RobustBoost*. Além disso, foi utilizado o sinal de ECG em uma técnica de aprendizagem profunda denominada Rede Neural Convolutiva (CNN, do inglês *Convolutional Neural Network*). E, por último, foram realizadas análises estatísticas a fim de comparar os resultados obtidos pelos diferentes classificadores.

Os resultados de medidas de desempenho dos sistemas classificadores indicaram que é possível classificar um sinal de ECG em duas classes distintas, diferenciando indivíduos com SM dos demais, com valores estatisticamente significativos. Utilizando características extraídas dos sinais de ECG, médias e variâncias das relações e intervalos entre os picos Q, R e S consecutivos e o eixo cardíaco, notou-se que SVM e *RobustBoost* obtiveram valores de exatidão média de aproximadamente 94% e 89%, respectivamente. Além disso, 12 derivações dos sinais de ECG foram utilizadas para o treinamento da CNN, que obteve acurácia de aproximadamente 98%. Portanto, conclui-se que é possível detectar a SM a partir dos sinais de Eletrocardiografia com resultados estatisticamente significativos, utilizando diferentes abordagens. Logo, um eletrocardiógrafo pode ser utilizado para avaliar um possível diagnóstico de SM.

ABSTRACT

The metabolic syndrome (MS) components is mainly associated with insulin resistance and diabetes. In addition, MS causes physiologic alterations that can induce injury to health and quality of patients lives. Scientific studies associate these changes with cardiovascular risk. These studies indicate that features extracted from electrocardiography (ECG) signals may be associated with MS. Despite this, scientific articles about MS detection with ECG signals were not found. Therefore, it is important to study elements of cardiovascular system that can indicate MS, in order to develop strategies for MS detection by means of ECG signals. Because of this scientific gap, the proposal of this work is to develop and to evaluate tools for MS diagnose with ECG signals.

To evaluate how accurately and precision classifier systems detect MS from ECG signals, the following procedures were adopted. Initially, Q, R and S peaks were detected from ECG signals waveform to extract temporal features. These features were used for training and validation of classifier systems, using two techniques, named Support Vector Machines (SVM) and RobustBoost. In addition, the ECG signal was used in a deep learning technique, named Convolutional Neural Network (CNN). Finally, statistical analyzes were performed in order to compare results of different classifiers.

The performance measures results of classifier systems indicated that it is possible to classify an ECG signal in two different classes, differentiating people with MS from others, with statistically significant values. With features extracted from ECG signals, means and variances of ratio and intervals between consecutive Q, R and S peaks and cardiac axis, it was noted that SVM and RobustBoost have average accuracy of approximately 94% and 89%, respectively. In addition, 12-lead ECG it was used in CNN to classify with approximately 98% of accuracy. Therefore, it is concluded that it is possible detect MS from electrocardiography signals with statistically significant results, using different approaches. Therefore, an electrocardiograph can be used to indicate a possible diagnosis of MS.

SUMÁRIO

Resumo	vi
<i>Abstract</i>	vii
1 Introdução	1
1.1 Objetivos	3
1.1.1 Objetivo Geral	3
1.1.2 Objetivos Específicos	4
2 Fundamentação teórica e o Estado da Arte	5
2.1 Sinais de Eletrocardiografia	5
2.1.1 Potenciais de Ação Gerados nas Membranas Plasmáticas que Resultam no Sinal de Eletrocardiografia	6
2.1.2 Formação e Aquisição dos Sinais de Eletrocardiografia	8
2.1.3 Diferentes derivações obtidas no plano frontal e horizontal	8
2.2 Extração de características dos Sinais de Eletrocardiografia	13
2.2.1 Algoritmo de Pan-Thompkins	13
2.3 Aprendizagem de máquina e sinais biológicos	14
2.3.1 Máquinas de Vetores de Suporte	15
2.3.2 <i>Ensembles</i>	18
2.3.3 Redes Neurais Convolucionais	19

2.3.4	Sinais de Eletrocardiografia e Aprendizagem de Máquina	21
2.4	Síndrome Metabólica e o Risco Cardíaco	22
2.4.1	Obesidade	23
2.4.2	Triglicerídeo	24
2.4.3	Colesterol	24
2.4.4	Pressão arterial	24
2.4.5	Nível de glicose sanguínea	24
3	Metodologia	27
3.1	Algoritmos desenvolvidos para classificação dos sinais de ECG	27
3.1.1	Extração de características dos sinais de eletrocardiografia	28
3.1.2	Desenvolvimento de sistema classificador	33
3.2	Metodologia experimental	34
3.2.1	Banco de dados	34
3.2.2	Procedimentos utilizados para desenvolvimento dos sistemas clas- sificadores	35
3.3	Análises estatísticas	37
3.3.1	Análises estatísticas dos resultados obtidos pelos algoritmos de de- tecção dos picos do sinal de ECG	38
3.3.2	Métricas de desempenho dos sistemas classificadores testados	39
3.3.3	Testes estatísticos de hipótese	41
4	Resultados e Discussão	43
4.1	Análises estatísticas dos resultados de detecção dos picos dos sinais de ECG	43
4.2	Métricas estatísticas de desempenho dos sistemas classificadores	45

4.2.1	Análise dos resultados obtidos pelos classificadores	50
4.3	Testes estatísticos de hipóteses	51
5	Conclusão	53
	Apêndice A	63
	Apêndice B	65
	Apêndice C	67

LISTA DE TABELAS

1.1	Fatores de risco da síndrome metabólica de acordo com o critério NCEP-ATP III. Adaptado de [1].	1
4.1	Média e desvio padrão das métricas de desempenho em porcentagem dos sistemas classificadores analisados em um teste com dados selecionados aleatoriamente, utilizando k-fold com k igual a 10, em apenas uma derivação dos sinais de ECG (derivação <i>DI</i>). Sendo <i>TFP</i> a taxa de falsos positivos (erro tipo I), <i>TFN</i> é a taxa de falsos negativos (erro tipo II), <i>VPP</i> é o valor preditivo positivo, <i>VPN</i> é o valor preditivo negativo, <i>S</i> é a sensibilidade, <i>Ex</i> é a exatidão, <i>Es</i> é a especificidade e <i>MF</i> é a medida-F. 47	47
4.2	Média e desvio padrão das métricas de desempenho em porcentagem dos sistemas classificadores analisados em um teste com 500 iterações de treinamentos e validações com dados selecionados aleatoriamente, utilizando k-fold com k igual a 10, em duas derivações dos sinais de ECG (derivação <i>DI</i> e <i>aVF</i>). Sendo <i>TFP</i> a taxa de falsos positivos (erro tipo I), <i>TFN</i> é a taxa de falsos negativos (erro tipo II), <i>VPP</i> é o valor preditivo positivo, <i>VPN</i> é o valor preditivo negativo, <i>S</i> é a sensibilidade, <i>Ex</i> é a exatidão, <i>Es</i> é a especificidade e <i>MF</i> é a medida-F.	48
4.3	Média e desvio padrão das métricas de desempenho em porcentagem dos sistemas classificadores analisados em um teste, utilizando o método k-fold, com $k = 10$, utilizando SVM, <i>RobustBoost</i> e CNN. Sendo <i>TFP</i> a taxa de falsos positivos (erro tipo I), <i>TFN</i> é a taxa de falsos negativos (erro tipo II), <i>VPP</i> é o valor preditivo positivo, <i>VPN</i> é o valor preditivo negativo, <i>S</i> é a sensibilidade, <i>Ex</i> é a exatidão, <i>Es</i> é a especificidade e <i>MF</i> é a medida-F. 48	48
4.4	Moda aproximada da exatidão obtida pelos diferentes classificadores em 10 testes realizados utilizando k-fold com k igual a 3 nos sinais de ECG adquiridos durante o estágio Basal.	49

4.5	Tempo aproximado de execução dos algoritmos, utilizando a técnica k-fold para separar a quantidade de dados para treinamento e validação, com k igual a 10, em um teste realizado com cada sistema classificador, utilizando o mesmo <i>hardware</i> e mesmo sistema operacional.	49
4.6	Valor p do resultado do teste de hipótese de cada característica extraída dos sinais de ECG, relacionados ao grupo controle e ao grupo com SM, dos diferentes estágios do TOTG.	52
C.1	Legenda das características extraídas dos sinais de ECG apresentadas na Tabela 4.6.	67

LISTA DE FIGURAS

2.1	Modelo de um sinal de Eletrocardiografia. Fonte própria.	5
2.2	Modelo do potencial de ação da célula nervosa. Adaptado de [23].	6
2.3	Modelo do potencial de ação da célula cardíaca. Adaptado de [26].	7
2.4	Trocas iônicas nas células cardíacas que ocorrem nas 5 fases durante a despolarização e a repolarização. Fonte própria.	7
2.5	Representação gráfica de um sinal de eletrocardiografia. O gráfico superior consiste em um modelo de um sinal de ECG indicando as fases de despolarização e repolarização e o gráfico inferior consiste em um sinal de ECG real – frequência de amostragem de 1 kHz. Fonte própria.	9
2.6	Visualização gráfica de cada derivação do eixo cardíaco, incluindo as derivações frontal e horizontal. Fonte: [29].	10
2.7	Triângulo de Einthoven e derivações no plano frontal. Adaptado de [27, 26].	10
2.8	Derivações no plano horizontal (unipolares). Fonte: [26].	11
2.9	Espectro da energia relativa do complexo QRS, ondas P e T, ruído muscular e artefatos de movimento baseados em uma média de 150 batimentos. Adaptado de [30].	14
2.10	Exemplo de dados bidimensionais de duas classes distintas. Fonte própria.	15
2.11	Exemplo de treinamento de dados bidimensionais de duas classes distintas utilizando SVM com kernel linear. Fonte própria.	16
2.12	A arquitetura comum do Ensemble. Baseado em [33].	18
2.13	Camadas de uma Rede Neural Convolucional, que é dividida em dois estágios: extração de características e classificação. Fonte própria.	19
2.14	Função ReLU, $\forall x > -100$ e $x < 100$. Fonte própria.	20

2.15	Relação entre os fatores de risco da Síndrome Metabólica e o risco cardíaco. Valores limites de cada um desses fatores podem ser observados na Tabela 1.1. Fonte própria.	22
2.16	Número de mortes por diabetes causadas por fatores ambientais, comportamentais e metabólicos no mundo em 2018. Incluindo pessoas de todas as idades e de ambos os sexos. Adaptado de [8].	25
3.1	Diagrama das etapas do processo realizado com os Sinais de Eletrocardiografia, incluindo as etapas de extração e classificação. Fonte própria. . . .	28
3.2	Fluxograma do algoritmo desenvolvido para detectar os picos Q, R e S do sinal de ECG utilizando o algoritmo de Pan e Thompkins desenvolvido por [61]. Fonte própria.	29
3.3	Exemplo da inversão no domínio do tempo realizada no sinal de ECG para detecção do pico da onda Q, utilizando quatro etapas diferentes, sendo Etapa 1 – o sinal de ECG original; Etapa 2 – o sinal invertido; Etapa 3 – o sinal de ECG invertido com os picos Q detectados; e Etapa 4: inversão realizada novamente, voltando ao sinal original com os picos Q detectados. Fonte própria.	31
3.4	Exemplo do sinal de Eletrocardiografia utilizado para extrair os picos Q, R e S manualmente e comparar com os resultados obtidos pelo algoritmo de detecção automática. Fonte própria.	32
3.5	Fluxograma das etapas da metodologia experimental de treinamento e validação dos classificadores em um exemplo do k-fold com k igual a 3. Baseado em [76].	36
3.6	Camadas da Rede Neural Convolutacional projetada para o treinamento e classificação dos sinais de ECG em dois grupos (grupo controle e grupo com Síndrome Metabólica). Fonte própria.	38
3.7	Etapas das análises estatísticas realizadas nas características extraídas dos sinais de ECG e nos resultados obtidos pelos sistemas classificadores. Fonte própria.	38
3.8	Exemplo do gráfico de Bland Altman, onde foi medida a taxa de pico de fluxo expiratório (PEF). Adaptado de [63].	40

3.9	Matriz de confusão desenvolvida para análise das métricas de desempenho dos sistemas classificadores nos sinais de ECG do banco de dados conforme Subseção 3.2.1. Sendo VP quantidade de verdadeiros positivos, FN quantidade de falsos negativos, FP quantidade de falsos positivos e VN quantidade de verdadeiros negativos. Fonte própria.	40
4.1	Média em função da diferença entre as localizações dos picos Q utilizando o algoritmo de detecção automático e o algoritmo de detecção manual. Fonte própria.	44
4.2	Média em função da diferença entre as localizações dos picos R utilizando o algoritmo de detecção automático e o algoritmo de detecção manual. Fonte própria.	44
4.3	Média em função da diferença entre as localizações dos picos S utilizando o algoritmo de detecção automático e o algoritmo de detecção manual. Fonte própria.	45
4.4	Exatidões em função das frequências relativas obtidas pela SVM e pelo <i>RobustBoost</i> na classificação dos sinais de ECG adquiridos durante o estágio Basal, utilizando k-fold com k igual a 10 em 500 testes. Fonte própria.	47
A.1	Representação gráfica de um sinal de eletrocardiografia indicando as fases de despolarização e repolarização. Fonte própria.	63
A.2	Representação gráfica de um sinal de eletrocardiografia indicando seus principais segmentos e intervalos. Adaptado de [10].	64
B.1	Representação gráfica para análise do eixo cardíaco com respeito as derivações <i>DI</i> , <i>DII</i> , <i>DIII</i> , <i>aVR</i> , <i>aVL</i> e <i>aVF</i> do plano frontal. Adaptado de [26, 27].	65
B.2	Representação das regiões consideradas de desvio do eixo e a região considerada como normal depois da identificação do eixo cardíaco. Adaptado de [10, 29].	66

LISTA DE SÍMBOLOS, NOMENCLATURAS E ABREVIACÕES

- CNN ou ConvNet – Rede Neural Convolutacional – do inglês, *Convolutional Neural Network*
- ECG – Sinais de Eletrocardiografia
- GBBANet – Grupo de Bioengenharia e Biofísica Aplicada – do espanhol, *Grupo de Bioingeniería y Biofísica Aplicada*
- GC – Grupo controle
- HDL – Lipoproteína de Alta Densidade – do inglês, *High Density Lipoproteins*
- IHME – Instituto de Métricas e Avaliação em Saúde – do inglês, *Institute for Health Metrics and Evaluation*
- IMC – Índice de Massa Corporal
- IQR – Intervalo Interquartil – do inglês, *Interquartile Range*
- MLP – Perceptron Multicamadas – do inglês, *Multilayer Perceptron*
- NCEP-ATP III – III Painel de Tratamento Adulto do Programa Nacional de Educação sobre Colesterol – do inglês, *National Cholesterol Education Program's Adult Treatment Panel III*
- ReLU – Unidade Linear Retificada – do inglês, *Rectified Linear Unit*
- RNAs – Redes Neurais Artificiais
- SM – Síndrome Metabólica
- SNRdB – Relação Sinal-Ruído – do inglês, *signal-to-noise ratio*, em Decibéis
- SVM – Máquinas de Vetores Suporte – do inglês, *Support Vector Machine*
- TOTG – Teste Oral de Tolerância a Glicose – do inglês, *Oral Glucose Test Tolerance*

1 INTRODUÇÃO

A Síndrome Metabólica (SM) consiste em um conjunto de fatores de riscos relacionados ao aumento da obesidade abdominal, à redução do nível de colesterol do tipo Lipoproteína de Alta Densidade (HDL, do inglês *High Density Lipoproteins*), no sangue, à hipertensão e ao nível elevado de glicose sanguínea associada à resistência à insulina [1]. Logo o diagnóstico é realizado por meio da análise clínica desse conjunto de fatores. A associação desses diversos fatores provoca alteração no metabolismo e, conseqüentemente, mudanças no funcionamento do sistema cardiovascular. Entretanto essas alterações não são específicas desta síndrome.

Neste trabalho foi adotado o critério estabelecido pelo III Painel de Tratamento Adulto do Programa Nacional de Educação sobre Colesterol (NCEP-ATP III, do inglês *National Cholesterol Education Program's Adult Treatment Panel III*), que afirma que um indivíduo é diagnosticado com SM se tiver no mínimo três fatores de riscos classificados de acordo com os níveis apresentados na Tabela 1.1 [1]. Esses fatores de risco estão

Tabela 1.1. Fatores de risco da síndrome metabólica de acordo com o critério NCEP-ATP III. Adaptado de [1].

Fatores de risco	Nível
Circunferência da cintura (Homem)	>102 cm
Circunferência da cintura (Mulher)	>88 cm
Triglicerídeo	≥ 150 mg/dL
Colesterol HDL (Homem)	<40 mg/dL
Colesterol HDL (Mulher)	<50 mg/dL
Pressão arterial	≥ 130 mmHg (sistólica) ou ≥ 85 mmHg (diastólica)
Glicemia em jejum	≥ 110 mg/dL

correlacionados entre si. O aumento da gordura corporal, por exemplo, está relacionado ao aumento do triglicerídeo [2].

Apesar de não existir consenso na definição de SM, a literatura registra que os fatores de risco estão relacionados à resistência insulínica e à diabetes ou pré-diabetes, visto que

pacientes com SM e sem diabetes podem adquirir diabetes futuramente. Além disso, doenças cardiovasculares podem ocorrer em indivíduos com SM [3, 4, 5, 6, 7].

O número de casos de mortes por diabetes tem aumentado, segundo o Instituto de Métricas e Avaliações de Saúde (IHME, do inglês *Institute for Health Metrics and Evaluation*), passando de 5^o lugar em 1990 para o 3^o lugar nas causas de morte em 2018 no Brasil. E as principais razões de mortes por diabetes no mundo têm sido relacionadas com fatores metabólicos, incluindo, além do alto nível de glicemia em jejum, o aumento do índice de massa corporal (IMC) [8].

Essas alterações metabólicas podem refletir em risco de doenças cardiovasculares que são as principais causas de morte no mundo [8]. Sendo assim, o estudo dos sinais de Eletrocardiografia (ECG) dos indivíduos com SM é relevante para obter informações desses sinais que podem estar relacionadas a alterações geradas pela presença da síndrome.

O avanço das tecnologias na saúde que utilizam aquisição de sinais, como por exemplo, o eletrocardiograma, contribui para análise do funcionamento do sistema cardíaco, inclusive em tempo real, e prevenção de doenças cardiovasculares [9]. As principais informações analisadas em sinais de ECG por cardiologistas que podem indicar a presença de distúrbios do sistema cardiovascular são obtidas a partir dos intervalos e segmentos que relacionam a onda P, o complexo QRS e a onda T dos sinais de ECG. Esses intervalos e segmentos são obtidos dos sinais de ECG, conforme explicado no Apêndice A. Além disso, é relevante avaliar as informações obtidas pelo eixo cardíaco (Apêndice B), sobretudo do QRS, que corresponde à direção do vetor resultante da atividade elétrica no momento da despolarização ventricular, e que pode indicar, por exemplo, hipertrofia ventricular [10].

Diante disso, o objetivo dos sistemas automáticos de detecção de patologias a partir da análise do sinal de ECG consiste em obter informações, sobretudo no domínio do tempo, dos sinais de ECG para a classificação em sinais de um grupo saudável ou de um grupo patológico. Em relação ao contexto desta pesquisa, essas tecnologias são ferramentas que detectam algumas anormalidades dos sinais de ECG que podem ocorrer devido à presença da SM [11].

A literatura apresenta estudos relacionados às alterações cardíacas e à presença da SM, avaliando sinais de ECG [12, 13, 14]. Isso é possível, pois a SM está relacionada ao risco cardiovascular [15]. Fatores de risco da SM podem estar associados, por exemplo, à insuficiência cardíaca e ao risco de morte cardíaca súbita [16, 17]. Diante disso, parâmetros extraídos dos sinais de ECG, tais como, informações relacionadas aos batimentos cardíacos, às ondas P, ao complexo QRS e onda T obtidas por análises dos sinais

de ECG (Apêndices A e B), podem ser associados à presença da SM [12, 13, 6].

Entretanto, não foram encontrados artigos científicos que mencionem abordagens de detecção de SM utilizando sinais de ECG. Diante dessa lacuna, a proposta deste trabalho consistiu em desenvolver uma ferramenta de possível diagnóstico de SM por meio de sinais de ECG, visto que os fatores de risco da SM estão relacionados a alterações no sistema cardiovascular. Logo, é relevante utilizar o sinal de ECG em diferentes estratégias para detectar a Síndrome Metabólica, com o intuito de investigar elementos relacionados ao sistema cardiovascular que indiquem quadros suspeitos de Síndrome Metabólica.

Além desses estudos que incluem a extração de características seguida de classificação, uma das aplicações do Estado da Arte consiste em uso de técnicas de aprendizagens profundas na classificação dos sinais de ECG para detecção de doenças cardiovasculares, sem a necessidade de extração de características previamente [18, 19, 20, 21, 22].

Diante desses fatores apresentados, este trabalho consistiu, em continuidade a estudos anteriores [12], extrair informações dos sinais relacionadas aos intervalos e as relações extraídas do picos Q, R e S consecutivos dos sinais de ECG, incluindo o eixo cardíaco, para classificação de sinais de ECG na detecção de SM, utilizando dois tipos de classificadores denominados Máquinas de Vetores Suporte (SVM, do inglês *Support Vector Machine*) e *RobustBoost*. Além disso, utilizar um tipo de classificador de aprendizagem profunda denominada Rede Neural Convolutiva (CNN, do inglês *Convolutional Neural Network*). Portanto, o intuito foi comparar os resultados obtidos pelos diferentes classificadores, utilizando técnicas de análises estatísticas, tais como, medidas de exatidão, especificidade, sensibilidade e análises de histogramas.

1.1 OBJETIVOS

1.1.1 Objetivo Geral

Implementar e comparar sistemas classificadores de sinais de Eletrocardiografia para detecção de Síndrome Metabólica, utilizando técnicas que necessitam de extração de informações dos sinais relacionadas a análises temporais e ao eixo cardíaco e uma abordagem utilizando aprendizagem profunda; com o intuito de responder com que exatidão e sensibilidade é possível detectar a Síndrome Metabólica com base em apenas sinais de Eletrocardiografia.

1.1.2 Objetivos Específicos

- 1 Implementar um algoritmo de detecção de picos das ondas Q, R e S dos sinais de Eletrocardiografia para extração de características que são integradas a um sistema detector de Síndrome Metabólica;
- 2 Desenvolver sistemas de detecção de Síndrome Metabólica por meio de características extraídas dos sinais de Eletrocardiografia utilizadas no treinamento de Máquinas de Vetores de Suporte e RobustBoost;
- 3 Desenvolver um sistema classificador de Síndrome Metabólica com base em Rede Neural Convolutiva aplicada a sinais de Eletrocardiografia;
- 4 Comparar os resultados obtidos pelos diferentes classificadores, utilizando métricas estatísticas de desempenho.

2 FUNDAMENTAÇÃO TEÓRICA E O ESTADO DA ARTE

2.1 SINAIS DE ELETROCARDIOGRAFIA

O sinal de Eletrocardiografia (ECG) é um sinal adquirido na superfície do corpo. Esse sinal medido consiste no somatório dos diversos potenciais bioelétricos, chamados de potenciais de ação, que se inicializam nas células nervosas. Esses potenciais de ação ocorrem devido ao fluxo de íons na membrana plasmática das células. Essas atividades dos potenciais de ação são transmitidas até a superfície do corpo e captadas por eletrodos posicionados em pontos específicos na superfície do corpo. E como resultado dessas medidas obtém-se o sinal de ECG que possui formato padrão mostrado na Figura 2.1.

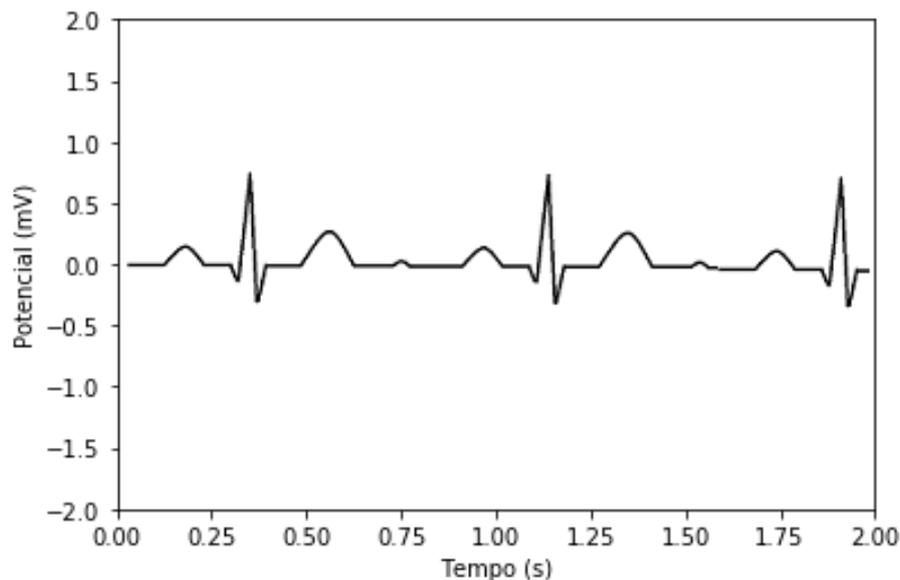


Figura 2.1. Modelo de um sinal de Eletrocardiografia. Fonte própria.

Esse formato possui relação com os potenciais de ativação das células musculares, potencial esse relacionado com a contração muscular. Como as células contraem normalmente em sequência padronizada, forma-se um sinal que, em pessoas saudáveis, em uma situação normal, tem um formato típico. Logo, a fim de entender a formação e o comportamento do sinal de ECG, é necessário compreender primeiramente o potencial

de ação que ocorre nas membranas plasmáticas das células.

2.1.1 Potenciais de Ação Gerados nas Membranas Plasmáticas que Resultam no Sinal de Eletrocardiografia

Ao se estudar a formação dos sinais de ECG, o primeiro passo consiste em compreender como ocorrem atividades elétricas que se inicializam nas células dos neurônios, devido trocas iônicas que ocorrem na membrana plasmática, e resultam nesses sinais [23].

Essas atividades iônicas nas células estão relacionadas a impulsos elétricos denominados potenciais de ação. O potencial de ação corresponde a alteração do potencial de membrana plasmática representada na Figura 2.2, que apresenta o potencial de ação de uma célula nervosa.

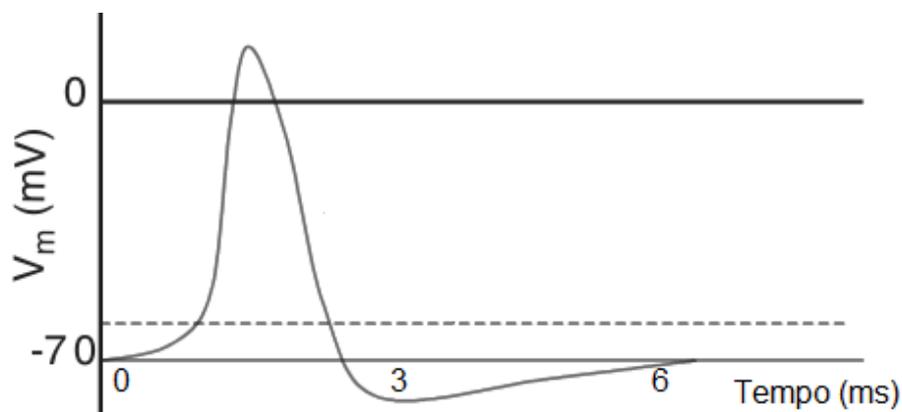


Figura 2.2. Modelo do potencial de ação da célula nervosa. Adaptado de [23].

No estado de repouso, o interior da célula encontra-se em um potencial negativo, que é em torno de -70 mV. A partir de um estímulo mínimo, quando é alcançado o potencial de limiar, ocorre a ativação de trocas iônicas. Em seguida, inicia-se a fase de despolarização, que corresponde à abertura dos canais de sódio (entrada de sódio na célula). Nesse momento o potencial no interior da célula aumenta, tornando-o positivo. Em seguida, ocorre a fase da repolarização, que corresponde ao fechamento dos canais de sódio e abertura dos canais de potássio (saída do potássio da célula). Devido ao desbalanceamento iônico (grande quantidade de sódio dentro da célula e de potássio fora da célula) é ativada a bomba de sódio e potássio que promove o mecanismo inverso (saída do sódio e entrada de potássio), fenômeno que ocorre com um gasto de energia gerado pela adenosina trifosfato (ATP). Neste estágio, a célula está mais eletronegativa do que era anteriormente ao estímulo. Por isso, ocorre o estágio de pós-hiperpolarização em que a célula retorna ao potencial de repouso [24, 25].

Nas células dos músculos cardíacos, este fenômeno de despolarização e repolarização ocorre em cinco fases, conforme apresentado na Figura 2.3: fase 0 (despolarização), fase 1 (repolarização precoce), fase 2 (repolarização lenta ou platô), fase 3 (repolarização rápida) e fase 4 (repouso ou polarização) [24].

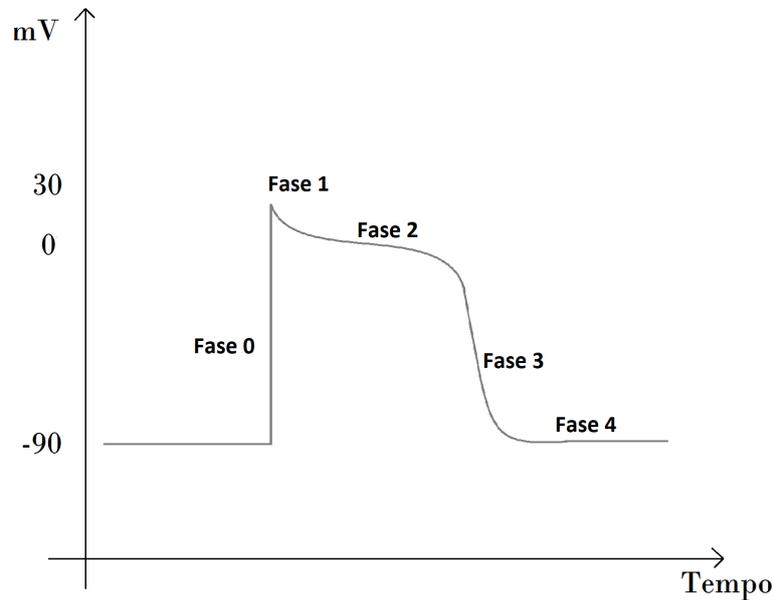


Figura 2.3. Modelo do potencial de ação da célula cardíaca. Adaptado de [26].

De forma similar ao modelo do potencial de ação da célula nervosa, em repouso o meio intracelular é negativo em relação ao meio extracelular, com distribuição de carga uniforme, ou seja, não existe diferença de potencial no interior e no exterior da célula. Entretanto, existe diferença de potencial entre o meio intra e o meio extracelular, denominado potencial de repouso, que é de aproximadamente -90mV. Quando a célula recebe um estímulo a nova diferença de potencial recebe o nome de potencial de ação [27].

As trocas iônicas que ocorrem nessas fases nas células cardíacas podem ser visualizadas na Figura 2.4. A partir deste estímulo, na fase 0 ocorre abertura dos canais de sódio

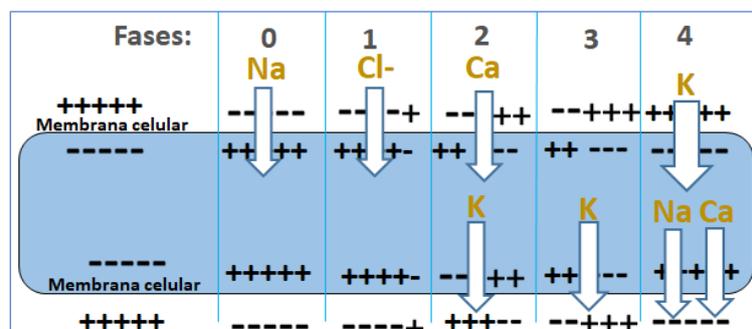


Figura 2.4. Trocas iônicas nas células cardíacas que ocorrem nas 5 fases durante a despolarização e a repolarização. Fonte própria.

(entrada de sódio) aumentando o potencial da membrana. Na fase 1, ocorre o fechamento dos canais de sódio e entrada de ânions cloro, reduzindo o potencial. Na fase 2, ocorre a abertura dos canais de cálcio (entrada de cálcio) e saída lenta de potássio, nesse momento a tensão elétrica não é alterada. Na fase 3, a redução do potencial ocorre devido a saída de potássio com rapidez, com o potencial retornando para $-90V$. Quando a saída de potássio excede a entrada de cálcio, ocorre a fase 4. Para compensar o excesso de sódio, a bomba de sódio e potássio elimina o sódio e deixa entrar potássio; além disso, o excesso de cálcio é eliminado. Nesta fase 4, finalmente o potencial da membrana permanece estabilizado até receber um novo estímulo e reiniciar o ciclo [26].

2.1.2 Formação e Aquisição dos Sinais de Eletrocardiografia

Essas atividades de trocas iônicas que ocorrem nas membranas celulares dos músculos cardíacos geram correntes elétricas que são transmitidas pelo organismo até a superfície do corpo. Essas correntes são medidas pelos eletrodos que são colocados em pontos específicos na superfície do corpo. O sinal de ECG corresponde a uma soma dos diversos potenciais de ação de várias células das regiões entre os eletrodos. A contribuição de cada célula depende de sua distância em relação ao eletrodo.

Um ciclo do sinal de ECG possui formato padrão conforme está apresentado na Figura 2.5. Ele é formado pela onda P, que representa o período de despolarização atrial, o complexo QRS, que representa a despolarização ventricular, e a onda T, que representa a repolarização ventricular. Em alguns casos, depois da onda T, ocorre um somatório de potenciais que formam a onda U.

2.1.3 Diferentes derivações obtidas no plano frontal e horizontal

Os sinais de ECG representam vetores cardíacos definidos por diferentes derivações. Um ECG padrão contém 12 derivações adquiridas a partir de 10 eletrodos posicionados em diferentes regiões. Esses eletrodos são colocados em posições próximas ao osso, evitando regiões musculares que geram sinais de eletromiografia e artefatos de movimentos que podem interferir nos sinais de ECG. Eles são nomeados conforme suas posições; logo as nomenclaturas usuais de cada eletrodo são *RA*: braço direito – do inglês, *Right Arm*; *LA*: braço esquerdo – do inglês, *Left Arm*; *RL*: perna direita – do inglês, *Right Leg*; *LL*: perna esquerda – do inglês, *Left Leg*; e *V1*, *V2*, *V3*, *V4*, *V5* e *V6* localizados na região da costela e mostrados na Figura 2.8 [28].

A partir dessas posições são obtidos sinais de ECG de dois diferentes tipos de de-

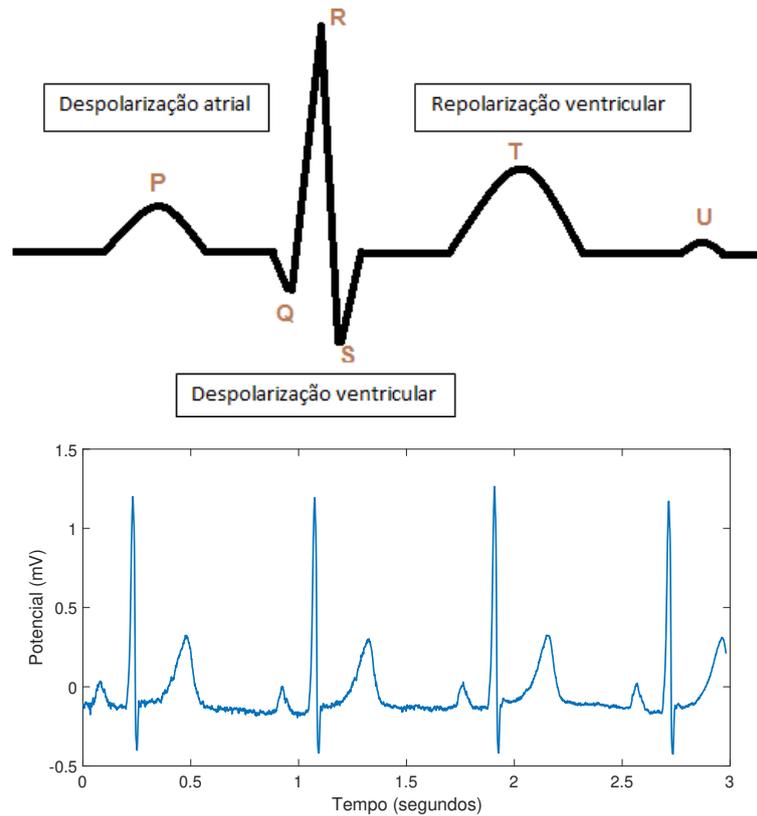


Figura 2.5. Representação gráfica de um sinal de eletrocardiografia. O gráfico superior consiste em um modelo de um sinal de ECG indicando as fases de despolarização e repolarização e o gráfico inferior consiste em um sinal de ECG real – frequência de amostragem de 1 kHz. Fonte própria.

rivações: as unipolares ou precordiais, registradas a partir de um ponto em relação a um ponto neutro, e as bipolares que são registradas como diferenças entre dois pontos [26]. Como o coração é uma estrutura tridimensional, mostrada na Figura 2.6, a análise de sua atividade elétrica ocorre também de forma tridimensional [10, 29]. Diante dessa estrutura do coração, as derivações foram definidas por dois planos, o plano frontal, que pode ser analisado segundo o eixo cardíaco, explicado na Subseção 2.1.3, e o plano horizontal [26].

Plano frontal As derivações foram inicialmente extraídas a partir do triângulo equilátero de Einthoven, conforme Figura 2.7 [30]. Esse triângulo originou-se como uma representação do posicionamento dos eletrodos, que estariam posicionados nos braços e na perna esquerda e no centro estaria o coração [26]. Portanto, o centro geométrico do triângulo de Einthoven – o ponto de intersecção das bissetrizes – representa o centro elétrico do coração, onde todos os vetores cardíacos possuem origem [27].

A partir de posições dos eletrodos relacionadas a pontos específicos na superfície do corpo, tais como, RA , LA , RL , e LL são encontradas diferentes derivações [28].

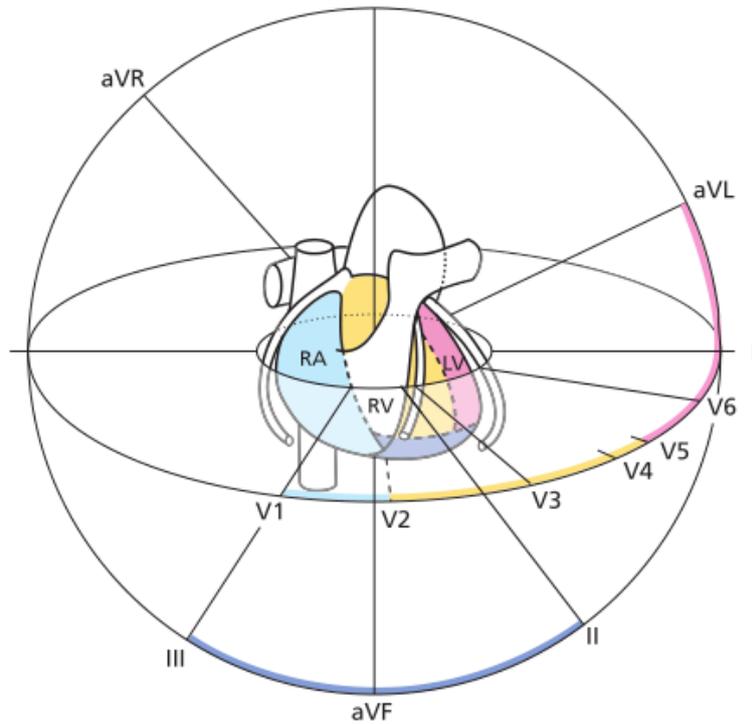


Figura 2.6. Visualização gráfica de cada derivação do eixo cardíaco, incluindo as derivações frontal e horizontal. Fonte: [29].

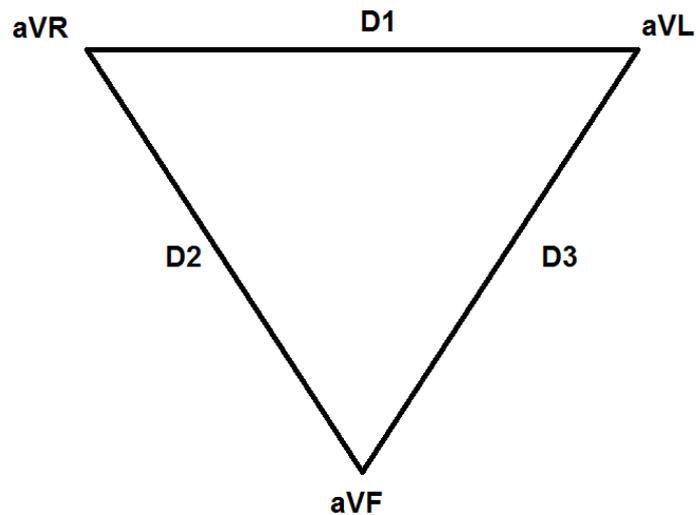


Figura 2.7. Triângulo de Einthoven e derivações no plano frontal. Adaptado de [27, 26].

Além disso, o ponto central, denominado terminal central de Wilson, é definido por [28]

$$V_w = \frac{1}{3(RA + LA + LL)}.$$

As derivações no plano frontal são extraídas do triângulo de Einthoven, conforme a Figura 2.7. Três são bipolares (DI , DII , $DIII$) [26]. Essas derivações são obtidas

por [28]

$$DI = LA - RA,$$

$$DII = LL - RA,$$

$$DIII = LL - LA.$$

Nessas três derivações (DI , DII , $DIII$), qualquer vetor no plano frontal pode ser projetado [27].

Além dessas derivações, três são unipolares, denominadas, aVL , aVF e aVR (sendo, a de aumentada – do inglês, *augmented*, V de eletrodo flutuante – do inglês, *roving electrode*, L , R e F de esquerda – *left*, direita – *right*, e membro inferior – *foot*, respectivamente) [26]. Essas derivações consistem em [28],

$$aVR = RA - \frac{1}{2(LA + LL)} = \frac{3}{2(RA - Vw)},$$

$$aVL = LA - \frac{1}{2(RA + LL)} = \frac{3}{2(LA - Vw)},$$

$$aVF = LL - \frac{1}{2(RA + LA)} = \frac{3}{2(LL - Vw)}.$$

Plano horizontal Além dessas, existem as derivações do plano horizontal que são unipolares e estão apresentadas na Figura 2.8. Essas derivações denominadas $V1$, $V2$, $V3$, $V4$, $V5$ e $V6$, estão localizadas na região da costela e mostrados na Figura 2.8 [26].

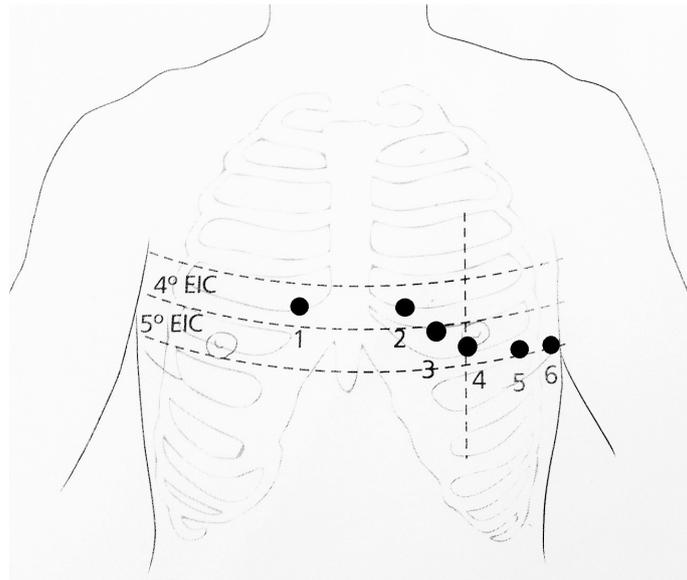


Figura 2.8. Derivações no plano horizontal (unipolares). Fonte: [26].

Eixo cardíaco As derivações no plano frontal (vertical), apresentadas na Figura 2.7, são derivações periféricas (*DI*, *DII*, *DIII*, *aVR*, *aVL*, *aVF*) [31]. Devido às rotações elétricas do coração que altera a relação dos vetores cardíacos com os eletrodos, essas derivações não são padronizadas [26], portanto são definidas segundo o eixo cardíaco. O eixo cardíaco consiste no vetor médio resultante dos potenciais de ação. Esse vetor corresponde ao direcionamento do fluxo elétrico, sobretudo durante o estágio de despolarização ventricular [29]. O eixo cardíaco é explicado em detalhes no Apêndice B.

Os desvios provocados no eixo podem indicar uma anormalidade. Geralmente ocorrem devido a uma hipertrofia no coração. Como o eixo dominante nas análises é o QRS, que representa a despolarização ventricular, um desvio no eixo indica, sobretudo uma hipertrofia ventricular [10]. Sendo assim, esse desvio no eixo pode indicar uma doença, ou ser uma variação normal que ocorrem, por exemplo, em crianças ou adultos magros. Um desvio pode indicar, por exemplo, um infarto do miocárdio, além disso, um desvio para a esquerda pode indicar hipertrofia ventricular esquerda e para a direita pode indicar uma hipertrofia ventricular direita [29].

Entretanto, a análise do eixo cardíaco não é suficiente para um diagnóstico, pois o coração realiza diversos deslocamentos. Esses deslocamentos ocorrem em relação a três eixos em diferentes sentidos – transversal ou látero-lateral, longitudinal e o antero-posterior. Nestes, as rotações ocorrem de ponta para frente e de ponta para trás; no sentido horário e no sentido anti-horário; e na posição horizontal, vertical e intermediária, respectivamente [31, 27].

Além disso, um mesmo coração possui combinações dessas rotações que podem ser encontradas tanto em cardiopatas como em indivíduos saudáveis. Portanto é importante analisar diversos fatores individuais relacionados às rotações elétricas de cada coração. Por exemplo, a rotação horária, posição vertical e ponta para trás pode ocorrer em pessoas longilíneas ou com crescimento do ventrículo esquerdo. Enquanto que a rotação anti-horária, posição horizontal e ponta para frente pode ser encontrada no coração de indivíduos brevilíneos ou com crescimento de ventrículo esquerdo [31].

2.2 EXTRAÇÃO DE CARACTERÍSTICAS DOS SINAIS DE ELETROCARDIOGRAFIA

2.2.1 Algoritmo de Pan-Thompkins

Pan e Thompkins desenvolveram em 1985 um algoritmo de extração do complexo QRS em tempo real dos sinais de ECG. Essa proposta tinha como objetivo a redução do ruído para detectar o complexo QRS, portanto possui um filtro digital passa faixa para reduzir interferências nos sinais [32].

Esse algoritmo utilizou três diferentes etapas, que incluíam os seguintes procedimentos:

- 1 – filtragem digital linear: filtro passa faixa, derivativo e integrador de janelas móveis;
- 2 – transformação não linear: quadratura de amplitude do sinal; e
- 3 – algoritmo de regras de decisão: limiares adaptativos e técnicas de discriminação da onda T [32].

O processo de filtragem envolve uma etapa de filtragem analógica, seguida de um conversor analógico digital e, por último, a filtragem digital [32].

O filtro passa faixas consiste em um filtro passa-baixa e um filtro passa-alta em cascata. Ele reduz as interferências de 60 Hz, as interferências dos movimentos, as linhas de bases e a interferência da onda T. Além disso, maximiza a energia do sinal entre 5 e 12 Hz [32]. Essa faixa corresponde a uma aproximação da faixa de frequência onde está localizado o espectro de energia do complexo QRS, conforme mostrado na Figura 2.9.

Portanto, após a digitalização do sinal, o algoritmo foi desenvolvido seguindo a seguinte sequência. Primeiramente, é projetado um filtro digital passa-faixas, em seguida, é realizada a diferenciação, a quadratura e o integrador de janelas móveis, respectivamente. Durante o estágio de derivação é obtida a inclinação do complexo QRS. Já a quadratura corresponde ao cálculo da energia do sinal. E finalmente, na etapa de regras de decisão é utilizada uma técnica de limiar dual, aplicada somente para sinais com batimentos cardíacos regulares. Nesta etapa, cada limiar é automaticamente calculado de acordo com o valor do pico do sinal e do ruído [32].

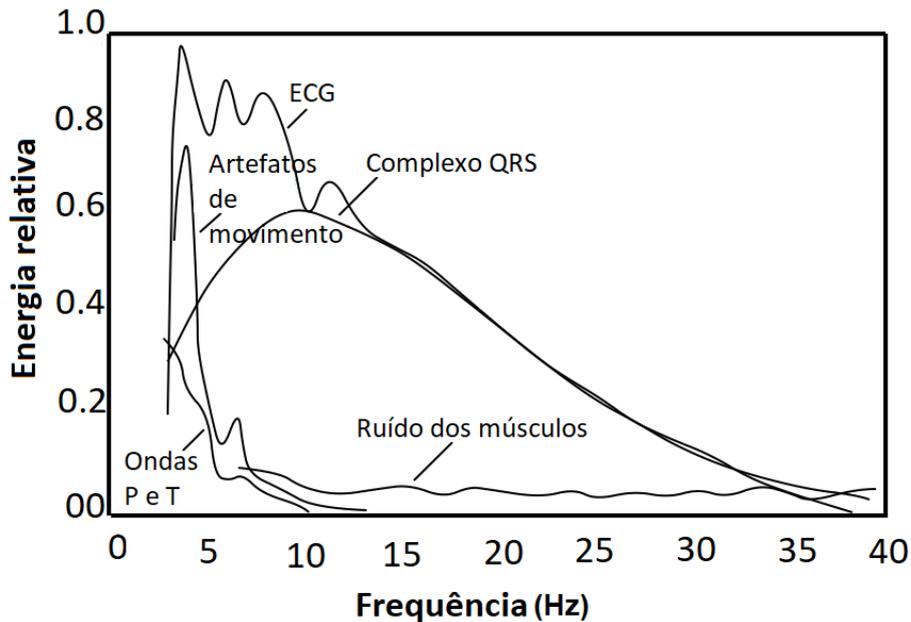


Figura 2.9. Espectro da energia relativa do complexo QRS, ondas P e T, ruído muscular e artefatos de movimento baseados em uma média de 150 batimentos. Adaptado de [30].

2.3 APRENDIZAGEM DE MÁQUINA E SINAIS BIOLÓGICOS

A aprendizagem de máquina consiste em técnicas de análises de dados, em que algoritmos são definidos para tomadas de decisões por meio do reconhecimento de padrões. Essas técnicas têm sido utilizadas na classificação de sinais. O intuito desta seção é detalhar o funcionamento dos principais classificadores utilizados em sinais biológicos segundo o Estado da Arte. A aplicação da aprendizagem de máquina em sinais de ECG está apresentada na Subseção 2.3.4.

A aprendizagem é dividida em aprendizagem supervisionada e aprendizagem não supervisionada. A aprendizagem supervisionada recebe os dados de treinamento rotulados, em outras palavras, recebe as classes dos dados. A aprendizagem não supervisionada não possui os dados rotulados, ela recebe apenas os dados e o sistema é treinado para identificar os diferentes agrupamentos.

Em um problema de classificação o rótulo do dado é categórico, como o caso deste trabalho, se o rótulo for numérico a análise seria realizada por métodos de regressão [33].

A aprendizagem consiste em duas etapas: no treinamento o sistema recebe os dados de entrada e define por meio de modelos matemáticos as diferenças entre as classes; e na validação ocorre a classificação dos dados e teste do sistema. Para validar um sistema são utilizadas algumas técnicas para estimação estatística, tais como validação cruzada e

bootstrap.

Na maioria dos classificadores binários, as classes são definidas como 1 e -1 (positivo e negativo). Diante disso, um classificador linear é um exemplo generalizado que é usado nos diferentes tipos de classificadores, e é modelado por

$$y = \text{sign}(w^T x + b),$$

onde w é o vetor de pesos, b é o viés, x é o vetor de instância e y é o resultado definido como a classe. Portanto, o sistema calcula a equação e, em seguida, define a região que separa os dados referentes às classes positivas e as classes negativas [33].

2.3.1 Máquinas de Vetores de Suporte

Máquinas de Vetores Suporte (SVM, do inglês *Support Vector Machine*), constituem um tipo de classificação supervisionada que utiliza otimização numérica para encontrar vetores de suporte. Esses permitem obter um hiperplano ótimo, em um espaço de dimensão maior, que separe as classes.

Um modelo que exemplifica a ideia básica da SVM é um dado bidimensional linearmente separável. A Figura 2.10 apresenta os dados de um exemplo de um sinal bidimensional, onde o eixo x representa uma característica e o eixo y representa a outra. Em verde estão representados os dados da classe 1 e em vermelho são os dados da classe -1.

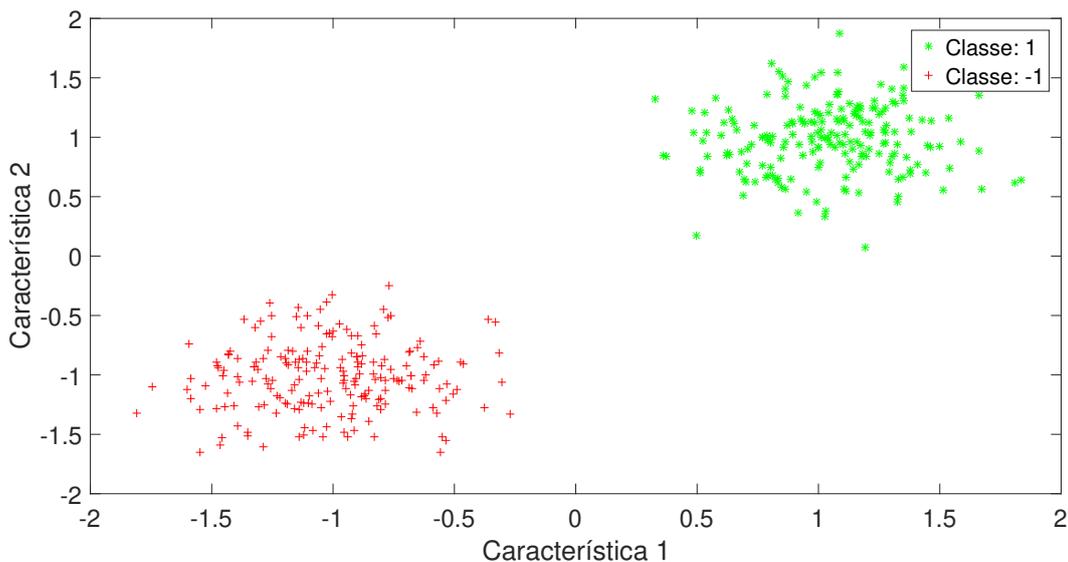


Figura 2.10. Exemplo de dados bidimensionais de duas classes distintas. Fonte própria.

Para encontrar o hiperplano que melhor separa essas duas classes, o sistema detecta os vetores de suporte quando recebe os dados de treinamento rotulados. Em seguida, é calculado o plano que maximiza a distância entre esses vetores de suporte, o hiperplano ótimo, que pode ser observado na Figura 2.11.

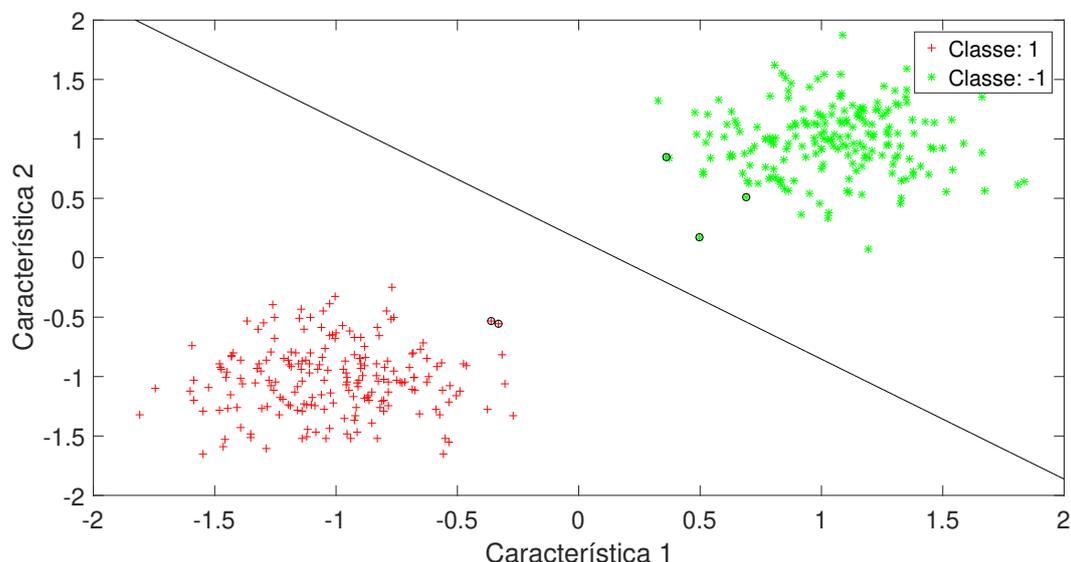


Figura 2.11. Exemplo de treinamento de dados bidimensionais de duas classes distintas utilizando SVM com kernel linear. Fonte própria.

O hiperplano ótimo é modelado por,

$$W^T X + b = 0, \quad (2.1)$$

onde W é o vetor de pesos que representa o direcionamento do hiperplano, o X é o vetor de características e b é um escalar que representa uma translação em relação a um plano que passa pela origem do sistema de coordenadas.

A partir da Equação 2.1, é possível distinguir duas classes. Para isso são usadas as equações

$$W^T X + b = 1, \quad (2.2)$$

e

$$W^T X + b = -1, \quad (2.3)$$

em que, caso o resultado da equação seja maior que zero o sistema classifica como uma classe (classe 1), se for menor que zero o sistema classifica como outra classe (classe -1), respectivamente.

A distância entre as Equações 2.2 e 2.3 consiste em

$$\frac{2}{\|w\|} = \frac{2}{\sqrt{w^T w}}. \quad (2.4)$$

O intuito é maximizar essa distância, e isso é semelhante a minimizar o inverso da Equação 2.4. Logo, temos o sistema de minimização de uma função, em que

$$\begin{aligned} \min \frac{w^T w}{2}, \\ \text{tal que } y_i((w^T x_i) + b) \geq 1, \end{aligned}$$

sendo $i = 1, \dots, L$.

Considerando o erro de classificação, temos

$$\begin{aligned} \min \frac{w^T w}{2} + C \sum_{i=1}^L \epsilon_i, \\ \text{tal que } y_i(w^T x_i) + b \geq 1 - \epsilon_i, \\ \epsilon \geq 0, \end{aligned}$$

sendo C uma constante e ϵ o erro dos *outlines* gerados na classificação.

Nos casos que não são linearmente separáveis a função dos dados é mapeada para um espaço de dimensão maior, que na maioria das vezes é um espaço de dimensão infinita, utilizando uma função ϕ associada ao *kernel* (núcleo) do sistema. Neste caso teremos,

$$\begin{aligned} \min \frac{w^T w}{2} + C \sum_{i=1}^L \epsilon_i, \\ \text{tal que } y_i(w^T \phi(x)_i) + b \geq 1 - \epsilon_i, \\ \epsilon \geq 0. \end{aligned}$$

A função de transformação ϕ se relaciona com o núcleo K pela equação

$$K(\mathbf{x}_i, \mathbf{x}_j) = \phi(x_i)^T \phi(x_j).$$

A Função de kernel mais utilizada é a função de base radial (do inglês, *Radial Basis Function*(*RBF*)), definida como

$$K(\mathbf{x}_i, \mathbf{x}_j) = e^{-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2}, \gamma > 0.$$

Além disso, são muito usados o núcleo linear

$$K(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^T \mathbf{x}_j,$$

a função polinomial,

$$K(\mathbf{x}_i, \mathbf{x}_j) = (\gamma \mathbf{x}_i^T \mathbf{x}_j + r)^d, \gamma > 0,$$

e a função Sigma,

$$K(\mathbf{x}_i, \mathbf{x}_j) = \tanh(\gamma \mathbf{x}_i^T \mathbf{x}_j + r),$$

onde γ , r e d são parâmetros das funções de cada núcleo [34].

2.3.2 Ensembles

O método de aprendizagem do tipo *Ensemble*, também chamado de aprendizagem baseada em comitês – do inglês, *committee-based learning*, constitui-se em um conjunto de diferentes classificadores utilizados em uma mesma classificação. Este tipo de aprendizagem é relevante, pois, geralmente, sua capacidade de generalização é maior do que a do classificador utilizado separadamente [33]. Uma arquitetura comum deste tipo de aprendizagem está representada na Figura 2.12.

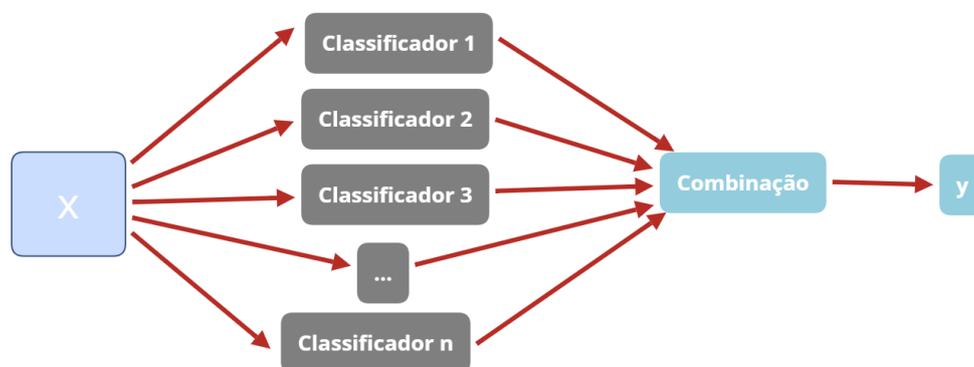


Figura 2.12. A arquitetura comum do Ensemble. Baseado em [33].

Existem dois tipos de aprendizagem de *Ensembles*: *ensemble* homogêneo, com aprendizagens do mesmo tipo e *ensemble* heterogêneo com diferentes tipos de aprendizagem [33].

Além dessa divisão base, existem três diferentes linhas nos métodos de aprendizagem por *ensemble*:

- 1 classificadores combinados: utilizada em reconhecimento de padrões; consiste em combinação de classificadores;

- 2 comitês de aprendizagem fraca: utilizados em aprendizagem de máquina; exemplos: *AdaBoost* e *Bagging*; e
- 3 mistura de especialidades: utilizado em redes neurais; aprendizagem com diversos modelos [33].

2.3.3 Redes Neurais Convolucionais

Em Aprendizagem Profunda – do inglês *Deep Learning*, as características não são mais extraídas por técnicas de processamento de sinais, são obtidas automaticamente durante o processo de aprendizagem. A Rede Neural Convolutiva (CNN, do inglês *Convolutional Neural Network*) consiste em um tipo específico de aprendizagem profunda, que contribuiu para o aperfeiçoamento na aprendizagem de máquina aplicada a imagens, vídeos, fala e áudio [35].

Alguns autores utilizaram as CNN nos sinais de ECG, em algumas aplicações tais como, detecção do complexo QRS [36] e classificação dos sinais de ECG com o intuito de detectar anormalidades cardíacas [18, 19, 20, 21, 22].

Sendo assim, para classificação dos sinais de ECG, diferentemente dos classificadores citados anteriormente, tais como, SVM e *Ensemble*, que realizam a classificação a partir de características previamente extraídas, a Rede Neural Convolutiva recebe como entrada o sinal completo e retorna como saída a classificação.

Para isso, é necessário o projeto da CNN, que inclui entre as camadas de entrada e de saída as camadas ocultas [35]. Dentre as camadas ocultas, existem dois estágios, mostrados na Figura 2.13, a extração de características e a classificação.

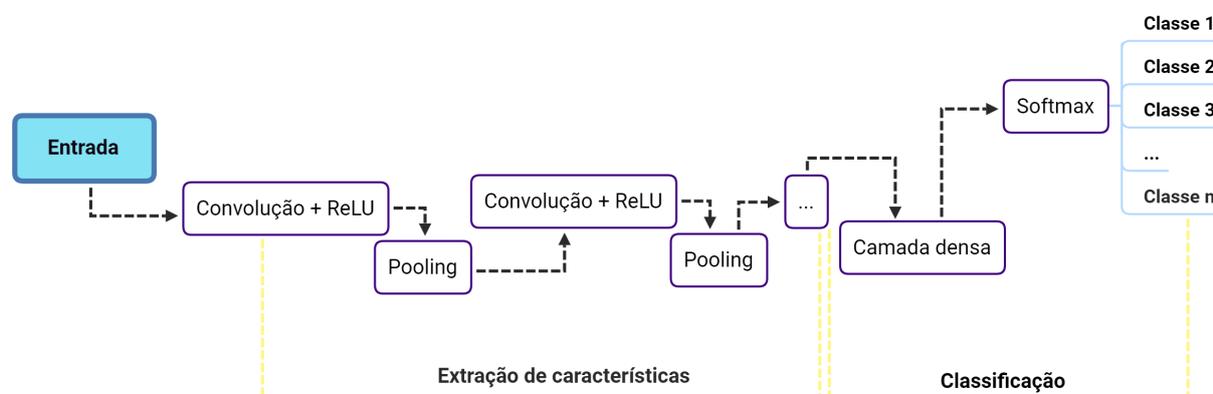


Figura 2.13. Camadas de uma Rede Neural Convolutiva, que é dividida em dois estágios: extração de características e classificação. Fonte própria.

No estágio da extração de características, o sinal passa pela etapa de convolução, pela Unidade Linear Retificada (ReLU, do inglês *Rectified Linear Unit*) e pela *maxpooling*. A etapa de convolução consiste em extrair características a partir de um processo de filtragem no sinal. A função ReLU é uma função não linear que consiste em zerar os valores negativos, sendo matematicamente representada por

$$f(x) = \begin{cases} 0 & \text{para } x < 0, \\ x & \text{para } x \geq 0, \end{cases}$$

onde x é a entrada da função e f é a função ReLU, tal que $f \in [0, \infty($. E graficamente representada pela Figura 2.14. E a etapa de *pooling* consiste em agrupar características

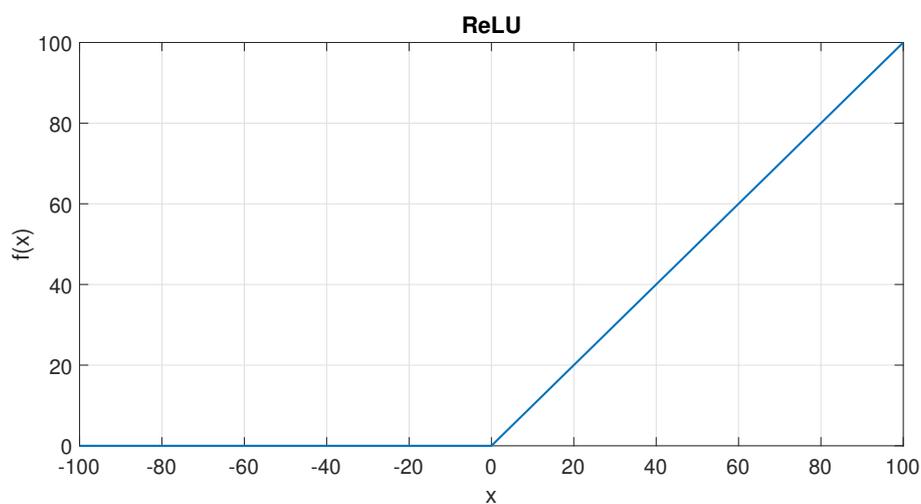


Figura 2.14. Função ReLU, $\forall x > -100$ e $x < 100$. Fonte própria.

semelhantes e, geralmente, extrair o máximo valor desse agrupamento, neste caso denominada *maxpooling* [35]. Estas etapas de convolução, ReLU e *maxpooling* são repetidas conforme a necessidade e especificação do projeto da *CNN*.

O estágio seguinte corresponde a classificação do sinal, o qual, já com as características extraídas, passa pela etapa de Camada Densa que determina a classe e pela função *Softmax*, matematicamente representada por

$$S(x_i) = \frac{\exp(x_i)}{\sum_{j=1}^k \exp(x_j)},$$

para $i = 1, \dots, k$, dado que x representa os valores de entrada da função e S consiste no resultado transformado. Essa função consiste em realizar cálculo de probabilidade transformando as classes de saída para valores entre 0 e 1.

2.3.4 Sinais de Eletrocardiografia e Aprendizagem de Máquina

A aprendizagem de máquina em sinais de ECG é utilizada para o desenvolvimento de sistemas automatizados de detecção de patologias. Ela consiste no uso de técnicas de reconhecimento de padrões para diferenciar de forma autônoma um sinal de ECG normal de um sinal com anormalidades provocadas por uma determinada doença. É importante ressaltar que para um diagnóstico um médico avalia, além dos resultados do eletrocardiograma, outras condições que identifique o paciente e podem alterar os sinais de eletrocardiografia, tais como, idade, histórico clínico e uso de medicamentos [27]. Portanto, o objetivo da aprendizagem de máquina no estudo dos sinais de ECG consiste em uma detecção de anormalidade no sinal que pode indicar a presença de uma patologia.

Existem diversos métodos de análises dos sinais de ECG. A literatura aponta estudos utilizando diferentes métodos de aprendizagem de máquina e de técnicas estatísticas e numéricas. Dentre elas, as técnicas mais utilizadas são as redes neurais artificiais (RNAs), as SVM, as análises de componentes principais e as transformadas de wavelet [28].

Foi realizado um estudo de comparação entre diferentes métodos, tais como, filtros, transformadas de wavelet, decomposições em modo empírico e redes neurais. Segundo este estudo, as redes neurais convolucionais possuem como vantagens a possibilidade de uso de diferentes tipos de características, alta acurácia, possibilidade de uso de técnicas estatísticas na seleção de características extraídas e análises de multi-resoluções. Entretanto, possui como desvantagens muito tempo e complexidade computacional [37].

O estado da arte trata do uso de técnicas de aprendizagens profundas, principalmente em estudos relacionados a processamento de imagens. Consequentemente, essa tendência tem sido aplicada na classificação dos sinais de eletrocardiografia no diagnóstico de doenças cardiovasculares, inclusive com o desenvolvimento de novos métodos de aprendizagem profunda [38]. Portanto, as pesquisas científicas apresentam avanços na área de aprendizagem profunda em análises de dados de sinais de eletrocardiografia [39].

Diversas pesquisas tratam de estudos relacionados às redes neurais e das CNN, em sinais de ECG [40], como por exemplo a localização de contração ventricular prematura utilizando 12 derivações de ECG [41] e a detecção da fibrilação atrial utilizando uma única derivação [42, 43], inclusive avaliando alguns segundos de sinais de ECG [44, 45, 46]. Em relação à acurácia dos sistemas classificadores, as redes neurais possuem resultados de acurácia estatisticamente significativos na classificação dos sinais de ECG [47]. Além disso, essa técnica torna-se importante, pois não há necessidade da extração de características previamente.

2.4 SÍNDROME METABÓLICA E O RISCO CARDÍACO

A Síndrome Metabólica (SM) consiste na associação de diversos fatores que geram alterações metabólicas que podem ocasionar doenças cardiovasculares. Ela possui como principal consequência a resistência a insulina, responsável pelo transporte e metabolismo da glicose. Logo os estudos que envolvem a SM consistem em relacionar a intolerância a glicose, e portanto, a predisposição da diabetes. A SM pode ser ocasionada por diversos fatores, tais como peso, gordura abdominal, diabetes e predisposição genética [2].

O III Painel de Tratamento Adulto do Programa Nacional de Educação sobre Colesterol (NCEP-ATP III, do inglês *National Cholesterol Education Program's Adult Treatment Panel III*) definiu que os fatores de risco da SM, apresentado na Tabela 1.1, são a circunferência da cintura acima de 102 cm nos homens e acima de 88 cm nas mulheres; taxa de triglicérideo maior ou igual a 150 mg/dL; colesterol do tipo Lipoproteína de Alta Densidade (HDL, do inglês *High Density Lipoproteins*), abaixo de 40 mg/dL nos homens e abaixo de 50 mg/dL nas mulheres; pressão arterial sistólica igual ou acima de 130 mmHg ou diastólica igual ou acima de 85 mmHg; e glicemia em jejum maior ou igual a 110 mg/dL. E para ser diagnosticado com SM é necessário no mínimo três desses fatores.

Devido a este conjunto de fatores, conforme diagrama mostrado na Figura 2.15, a SM está relacionada a consequências no sistema cardiovascular [48]. Diante disso, a seguir, são definidas as alterações fisiológicas e metabólicas causadas por cada um desses fatores e as consequências geradas no organismo humano relacionadas com o risco cardíaco [1].

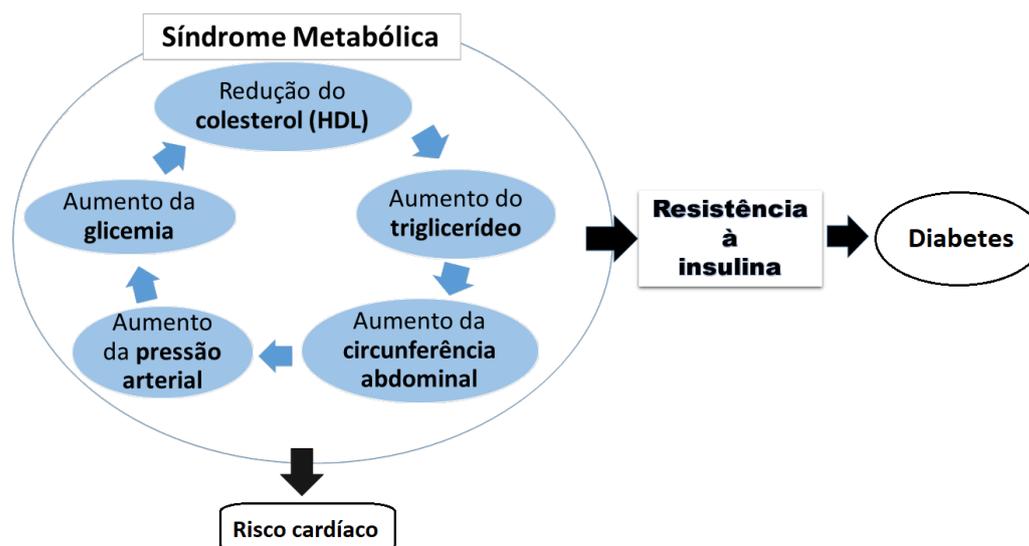


Figura 2.15. Relação entre os fatores de risco da Síndrome Metabólica e o risco cardíaco. Valores limites de cada um desses fatores podem ser observados na Tabela 1.1. Fonte própria.

Sendo assim, nesta seção são apresentadas relações entre os fatores de risco da SM e alterações provocadas nos sinais de ECG. Informações, tais como, complexo QRS, intervalo PR, onda P, onda T, intervalo QTc e eixo, citadas nesta seção, consistem nas principais características extraídas dos sinais de ECG e relevantes na análise e no diagnóstico clínico. Os Apêndices A e B explicam em detalhes o significado de cada uma dessas características.

2.4.1 Obesidade

Nem todos os indivíduos com SM são obesos, entretanto a obesidade é um dos principais fatores da SM que pode ocasionar alterações cardiovasculares [49]. Um exemplo de alteração cardiovascular é a presença da fibrilação atrial que pode ocorrer em indivíduos obesos e em indivíduos com SM [50].

Sendo assim, é relevante avaliar a relação de alterações metabólicas provocadas devido a obesidade [49]. Visto que, a gordura abdominal está intimamente relacionada com diversos fatores metabólicos, tais como resistência insulínica, aumento do triglicerídeo, redução do colesterol do tipo HDL e aumento do risco de diabetes [2].

Além disso, a obesidade é um dos fatores relevantes neste estudo, pois um indivíduo obeso, mesmo sem doença cardiovascular, possui maior tendência de ter arritmias. Isso porque a atividade simpática em obesos é maior e, conseqüentemente, ocorre risco de arritmias [11].

Não é recente a ideia de que o aumento da obesidade em um indivíduo pode ocasionar diversas alterações nos sinais de ECG. Em 1986, a obesidade foi correlacionada com o aumento da frequência cardíaca, e com parâmetros extraídos dos sinais de ECG, tais como, aumento do intervalo PR, da duração e da amplitude do complexo QRS, deslocamento para esquerda do complexo QRS, devido à redução da condução do sinal de ECG relacionado ao excesso de peso, e com o aumento da amplitude e do intervalo QTc (Apêndice A) [51].

Além disso, outro estudo de 2013 que avaliava os parâmetros alterados em crianças e adolescentes obesos foi observado aumento dos intervalos PR, da duração do complexo QRS, deslocamento para esquerda dos eixos frontais da onda P, do QRS e do eixo da onda T e aumento da frequência cardíaca. Além disso, a obesidade abdominal foi relacionada com o aumento do intervalo PR, aumento da duração do complexo QRS e deslocamento para esquerda do eixo frontal do ECG [52].

2.4.2 Triglicerídeo

O aumento dos níveis de triglicerídeo, sobretudo entre os indivíduos do sexo masculino [5], é outro fator de risco da SM para valores iguais ou acima de 150 mg/dL. Assim como os demais fatores, o triglicerídeo é analisado juntamente com outros fatores de risco da SM. O aumento do nível de insulina provoca acúmulo de triglicerídeo na região abdominal e, conseqüentemente, ganho de peso [53].

2.4.3 Colesterol

Outro fator de risco da SM é o colesterol do tipo HDL, também conhecido como *colesterol bom*. Por isso, ele é considerado um fator de risco da SM se estiver em um nível abaixo de 40 mg/dL nos homens e abaixo de 50 mg/dL nas mulheres. Em nível baixo, pode provocar risco de doença cardiovascular, como, por exemplo, doenças arteriais coronárias [54].

2.4.4 Pressão arterial

Assim como os demais fatores de risco da SM, a pressão arterial, associada a outros fatores, pode contribuir para ocorrência de doenças cardiovasculares. A hipertensão junto com a diabetes dobra o risco de doença cardiovascular [2]. E quando está associada com o excesso de peso pode ocasionar fibrilação atrial [50, 55].

2.4.5 Nível de glicose sanguínea

O nível de glicose está relacionado a capacidade da insulina em promover a entrada de glicose nas células [2]. Ele é o principal fator analisado em indivíduo com SM, relacionado à presença da pré-diabetes e da diabetes em que ocorrem quadros de hipoglicemia e hiperglicemia, podendo ocasionar, como consequência, risco cardíaco.

A diabetes ou a pré-diabetes são as principais causas de alterações cardiovasculares [56]. Sendo assim, estudos relacionados à diabetes apontam que é possível detectar as complicações da diabetes de uma forma não invasiva com acurácia acima de 95%, avaliando, por exemplo, a variabilidade da frequência cardíaca [57, 58].

Além disso, a diabetes está diretamente relacionada a alterações metabólicas. Segundo o Instituto de Métricas e Avaliação em Saúde (IHME, do inglês *Institute for*

Health Metrics and Evaluation), em 2018, as principais causas de morte por diabetes, no mundo, ocorreram por fatores metabólicos, a glicemia alta em jejum e altos índices de massa corporal (IMC), em comparação com os fatores ambientais (poluição do ar) e comportamentais – tais como, dietas de risco, fumo e baixa atividade física [8].

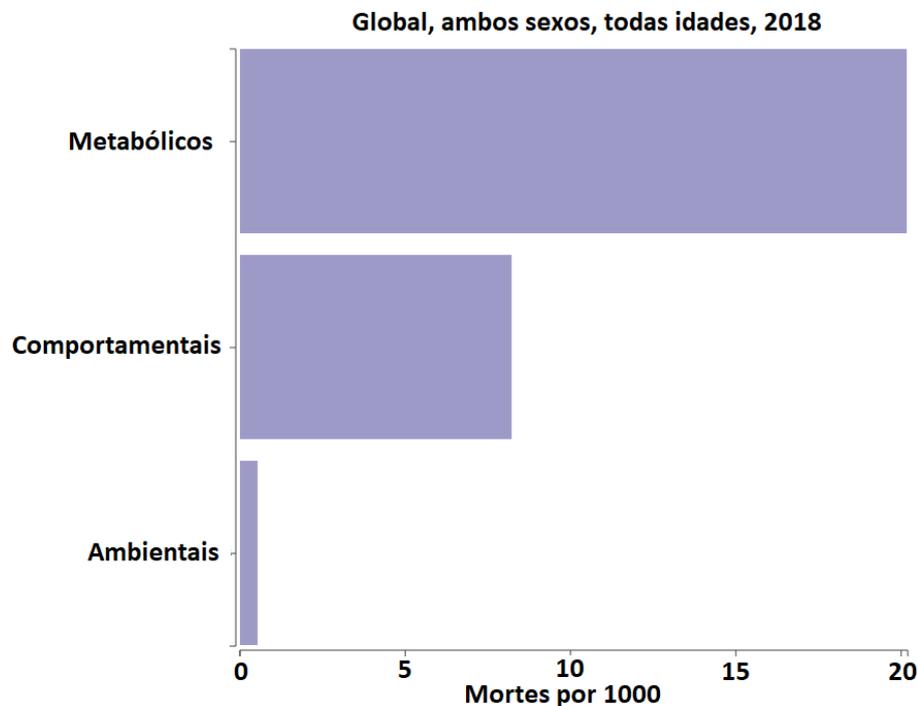


Figura 2.16. Número de mortes por diabetes causadas por fatores ambientais, comportamentais e metabólicos no mundo em 2018. Incluindo pessoas de todas as idades e de ambos os sexos. Adaptado de [8].

Apesar da influência da diabetes com os diversos fatores de risco da SM e a relação desses com alterações no sistema cardiovascular, um estudo apontou que pessoas com SM e sem diabetes não possuem aumento de risco cardíaco significativo em comparação com os indivíduos que possuem somente diabetes [59].

Indivíduos com diabetes podem passar por quadros de hiperglicemia e hipoglicemia. A hiperglicemia consiste em um drástico aumento de glicose no sangue que pode ocasionar aumento da frequência cardíaca [6]. Enquanto que, a hipoglicemia está relacionada à redução do nível de glicose sanguínea. Este quadro pode ocorrer devido efeito colateral da injeção de insulina. Em alguns casos, pode ocorrer, inclusive, durante o período noturno. Desta forma, um quadro de hipoglicemia ocasiona alterações no funcionamento do sistema cardiovascular [24].

Em um estudo de análises dos sinais de ECG de indivíduos em um momento de hipoglicemia induzida a fim de avaliar as variações ocorridas nesses sinais foram detectadas alterações em alguns parâmetros dos sinais de ECG (Apêndice A), como por exemplo,

deformidades na onda T, tais como largura da onda T, achatamento na onda, prolongamento dos segmentos QT e QTc ocorrem durante a hipoglicemia, indicando, portanto, que a hipoglicemia gera alterações durante o período de repolarização ventricular do sistema cardiovascular [24]. Além disso, em um estudo que utilizou sistemas classificadores de aprendizagem profunda, observou-se que é possível detectar um episódio de hipoglicemia em pacientes com diabetes, com sensibilidade acima de 80%, utilizando os intervalos QT e QTc e o batimento cardíaco [60].

3 METODOLOGIA

Esta pesquisa consistiu no desenvolvimento de diferentes abordagens para detecção da Síndrome Metabólica (SM) por meio de sinais de eletrocardiografia (ECG). Para isso, foram utilizadas estratégias distintas, detalhadas na Seção 3.1, para o treinamento e validação de sistemas classificadores de sinais de ECG. Os sinais utilizados nesta proposta são relacionados a dois grupos experimentais distintos, um grupo controle e um grupo com SM. Esses sinais são disponibilizados publicamente em um banco de dados apresentado na Subseção 3.2.1. Ao final foram realizadas medidas estatísticas, explicadas na Subseção 3.3.2, com intuito de avaliar com que exatidão e sensibilidade é possível detectar a SM por meio de sinais de ECG.

3.1 ALGORITMOS DESENVOLVIDOS PARA CLASSIFICAÇÃO DOS SINAIS DE ECG

Com o intuito de definir se o sinal de ECG possui um padrão considerado normal, ou se tem um desvio associado a Síndrome Metabólica (SM), a metodologia deste trabalho consistiu em duas etapas, uma de extração de características e outra de classificação do sinal. Para classificar, uma parte do sinal é utilizada para treinamento e outra para validação, que tem como objetivo verificar medidas estatísticas, apresentadas na Subseção 3.3.2, tais como, a eficácia e a exatidão do sistema após o treinamento.

O diagrama da Figura 3.1 apresenta as etapas necessárias para o desenvolvimento do sistema classificador do sinal de ECG. Para obter as características dos sinais que foram utilizadas nos sistemas classificadores, o sinal passou por um sistema detector de picos, em seguida foram extraídas características desses sinais, tais como, medidas de amplitude relativas e distâncias entre dois picos consecutivos. O intuito foi verificar se existem valores alterados no ECG e que possam estar associados a SM. Diante disso, o sistema classificador é um algoritmo que é ajustado e treinado para definir se um conjunto de características faz parte ou não de um exemplo de SM. Neste trabalho foram utilizados classificadores supervisionados, que possuem os rótulos definidos previamente, ou seja, as

classes do sinal (indivíduo saudável ou com SM) foram incluídas no treinamento.

Sendo assim, em resumo, primeiro foi realizada a extração de características dos sinais de ECG que foram, em seguida, utilizadas em sistemas que classificaram os sinais em dois grupos distintos – pacientes saudáveis e pacientes com SM.

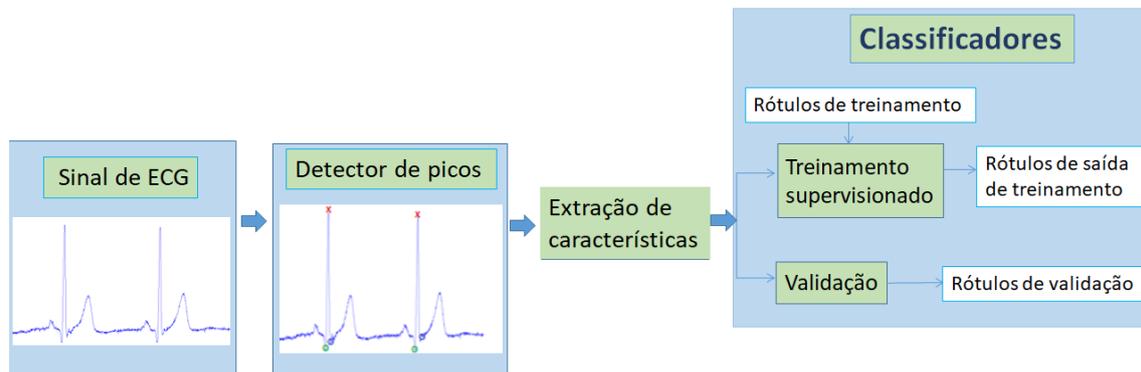


Figura 3.1. Diagrama das etapas do processo realizado com os Sinais de Eletrocardiografia, incluindo as etapas de extração e classificação. Fonte própria.

3.1.1 Extração de características dos sinais de eletrocardiografia

Detecção do complexo QRS Para extrair o complexo QRS foi utilizado um algoritmo, que implementa o algoritmo de Pan Thompkins para detecção do pico R do sinal de ECG, desenvolvido no Matlab por uma pesquisa realizada no Instituto de Pesquisa do Hospital Geral de Rochester, Rochester, Estados Unidos – do inglês, Rochester General Hospital Research Institute, Rochester, New York, United States [61].

Conforme explicado na Subseção 2.2.1, o algoritmo de Pan-Thompkins foi desenvolvido em 1985 e consiste em detectar o complexo QRS do sinal de ECG em tempo real [32]. Ele é apresentado na literatura como uma das técnicas de fácil implementação para a detecção do complexo QRS com o objetivo de extrair características dos sinais de ECG a serem utilizados em sistemas classificadores. Além disso, é uma técnica relevante, pois exige menos esforço computacional em comparação a outras técnicas e é projetado para reduzir o ruído [28, 47, 62].

Em um teste realizado utilizando esse algoritmo, observou-se que existem diferenças entre alguns pontos marcados como pico R e os picos máximos das ondas, portanto foi desenvolvido um algoritmo para ajuste de detecção e, também, de detecção dos picos das ondas Q e S. O fluxograma da Figura 3.2 apresenta as etapas para o desenvolvimento desta função. O ajuste de detecção dos picos R consiste em utilizar os dados dos picos R

extraídos e a partir de uma janela que avaliava valores próximos do pico detectado, foi realizada uma busca do pico máximo. Este valor foi então substituído como o novo valor de R.

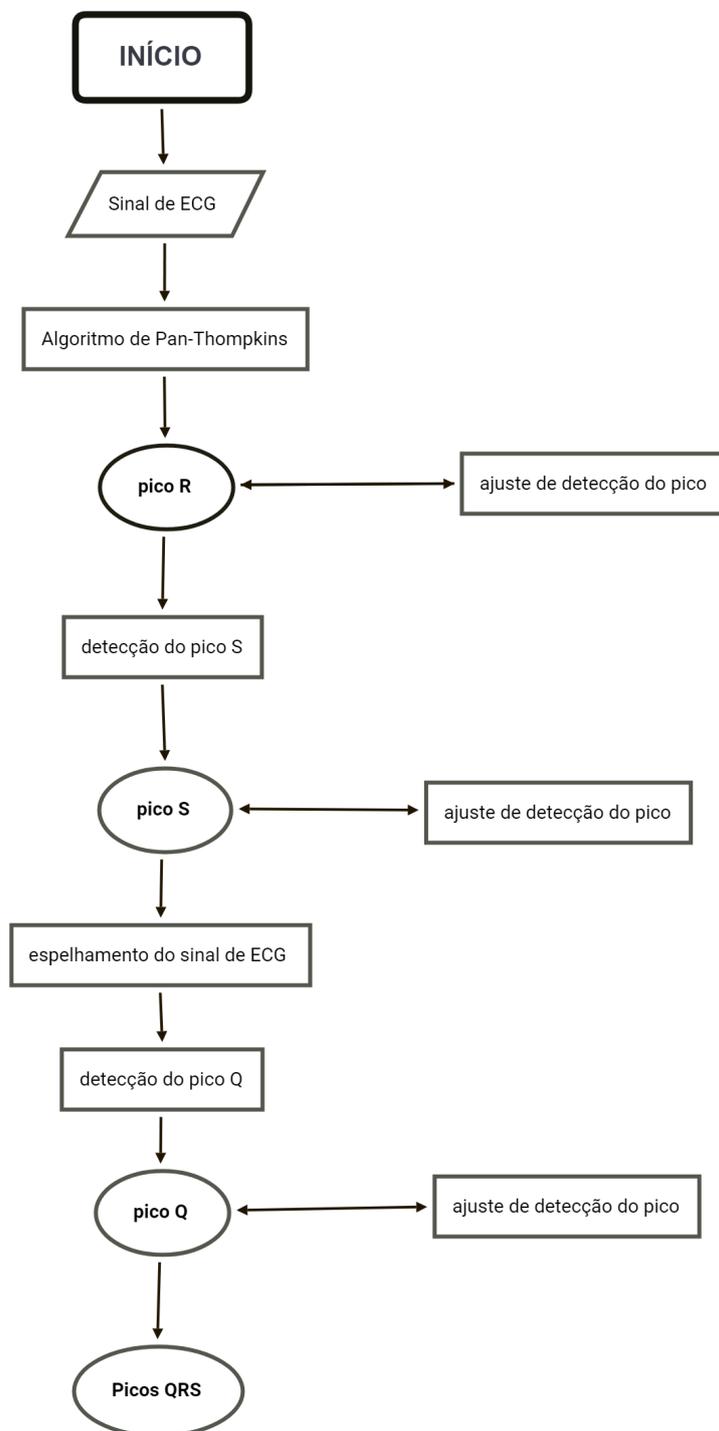


Figura 3.2. Fluxograma do algoritmo desenvolvido para detectar os picos Q, R e S do sinal de ECG utilizando o algoritmo de Pan e Thompkins desenvolvido por [61]. Fonte própria.

Em seguida, foi desenvolvido um algoritmo para detectar o pico S. Essa função consistiu em realizar cálculo de derivada para encontrar o ponto de mínimo da função mais próximo do pico R. Esse valor foi reajustado da mesma forma como ocorreu com o pico R.

Para encontrar o pico da onda Q do sinal de ECG foi realizado um espelhamento do sinal no sentido horizontal, ou seja, inversão do sinal no domínio do tempo, e então a localidade do pico Q assume a posição anterior do pico S, logo foi realizado o mesmo procedimento utilizado para detectar o pico S. Esse procedimento de detectar o pico Q, conforme mostrado na Figura 3.3, foi realizado nas quatro etapas seguintes:

- 1) o algoritmo recebe o sinal original com os picos R detectados;
- 2) aplicando a equação

$$x_e[n] = x[N - n],$$

esse sinal é invertido em relação aos valores no tempo;

- 3) mesmo procedimento utilizado para detectar o pico S, considerando também o ajuste; e
- 4) o sinal é invertido novamente voltando ao sinal original com os picos Q detectados.

Os valores das posições dos picos do complexo QRS – pico Q, R e S , conforme mostrado na Figura 2.5, dos sinais de ECG do grupo controle e do grupo experimental foram detectados separadamente e armazenados.

Validação do algoritmo detector de picos Com o intuito de verificar se os resultados de detecção dos picos podem interferir nas detecções de SM, ou seja, verificar a eficácia da extração de características dos sinais de ECG e sua influência na significância dos resultados obtidos pelos sistemas classificadores desenvolvidos, foi desenvolvido um algoritmo para detecção manual dos picos das ondas Q, R e S dos sinais de ECG. Desta forma, foi possível verificar o erro do algoritmo de detecção automática comparando com valores dos picos extraídos manualmente.

Esses resultados foram avaliados segundo o gráfico de Bland Altman, explicado na Subseção 3.3.1, que avalia se as diferenças obtidas por duas técnicas distintas são significativas e se seus resultados podem interferir em decisões aplicadas em diagnóstico [63].

Para detectar os picos foi desenvolvido no matlab um algoritmo de detecção manual, sendo realizado o mesmo procedimento para a detecção dos picos Q, R e S. O sinal

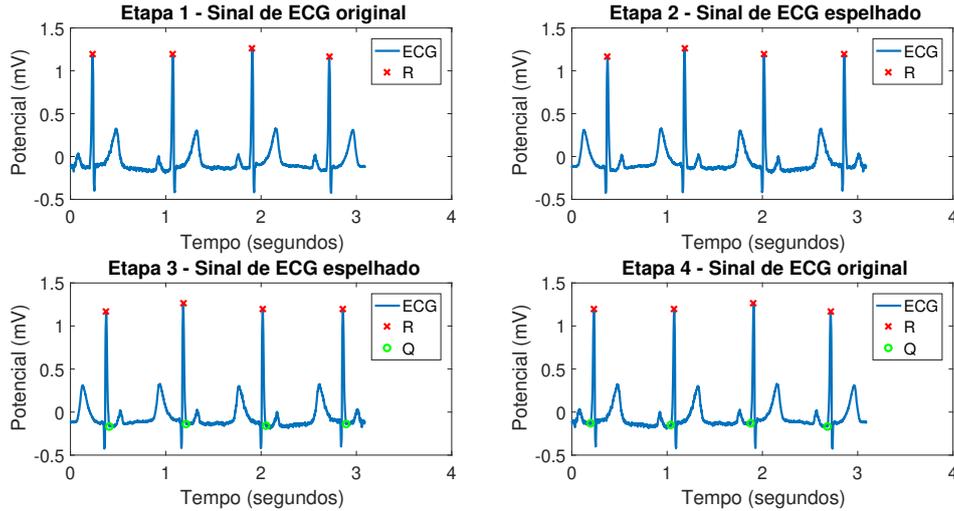


Figura 3.3. Exemplo da inversão no domínio do tempo realizada no sinal de ECG para detecção do pico da onda Q, utilizando quatro etapas diferentes, sendo Etapa 1 – o sinal de ECG original; Etapa 2 – o sinal invertido; Etapa 3 – o sinal de ECG invertido com os picos Q detectados; e Etapa 4: inversão realizada novamente, voltando ao sinal original com os picos Q detectados. Fonte própria.

completo possui duração de aproximadamente 30 minutos, para facilitar a visualização do sinal durante a detecção dos picos, ele foi dividido em janelas de aproximadamente 3 segundos, resultando em um total de 649 janelas.

Ao selecionar manualmente os valores dos picos, as localizações foram armazenadas em variáveis no algoritmo. Com duas seleções na mesma localização foi traçado o gráfico da região. Após a detecção do sinal completo, esta etapa foi finalizada e as localizações dos picos detectados foram armazenadas.

Um exemplo do sistema desenvolvido de detecção manual de pico para a onda R do sinal de ECG é mostrado na Figura 3.4.

Características referentes às distâncias e às relações entre Q, R e S A partir desses valores de QRS encontrados foi desenvolvido um algoritmo para extração de características no domínio do tempo. Essa função recebe como entrada o tamanho da janela e o arquivo com as localizações dos picos Q, R e S dos sinais de ECG detectados. Foram calculadas as médias e as variâncias dos intervalos RR, RS, QS, RQ, RS, da amplitude da onda R e das relações entre as amplitudes da onda R e da onda Q e da onda R e S, considerando a sobreposição do sinal, em janelas de tamanho 80. Ao final, foi gerada uma matriz com essas características extraídas. Essa matriz foi armazenada e posteriormente utilizada no treinamento e validação dos sistemas classificadores.



Figura 3.4. Exemplo do sinal de Eletrocardiografia utilizado para extrair os picos Q, R e S manualmente e comparar com os resultados obtidos pelo algoritmo de detecção automática. Fonte própria.

Em um estudo anterior foram avaliadas as características temporais de relações entre os picos dos sinais de ECG para o desenvolvimento de sistemas classificadores de SM, com resultados estatisticamente significativos, exatidão acima de 90%. Entretanto, foram avaliados apenas os sinais de ECG da derivação *DI* [12]. Na proposta dessa dissertação, incluiu-se a derivação *aVF*, além do eixo cardíaco. Portanto, foram extraídas características dos sinais de ECG referentes a duas derivações ortogonais do plano frontal, derivação *DI* e *aVF*. Esse método foi utilizado, pois essas derivações são analisadas em uma técnica de detecção do eixo cardíaco com melhor exatidão em relação a outras técnicas [64].

Extração do eixo cardíaco Como o coração é uma estrutura tridimensional, a análise de apenas uma derivação nem sempre é suficiente para indicar a existência de alguma anormalidade do sistema cardiovascular. Portanto, uma das estratégias para avaliar um desvio no sinal de ECG é realizar a análise do eixo cardíaco.

Estudos apresentam algumas alterações cardiovasculares ocasionadas em indivíduos com SM, como por exemplo, hipertrofia ventricular esquerda e comprometimento na função diastólica [65, 66]. O desvio de eixo cardíaco pode indicar a presença de uma alteração cardiovascular. Portanto foi proposto neste trabalho incluir o eixo cardíaco como uma das características extraídas dos sinais de ECG a serem utilizadas em sistemas de detecção de SM.

Para calcular o eixo cardíaco, são necessárias apenas duas derivações, conforme mostrado na Figura B.2. Logo, foram selecionadas as derivações *DI* e *aVF*, pois, conforme

o triângulo de Einthoven da Figura 2.7, são derivações ortogonais, logo o eixo cardíaco pode ser encontrado por meio de análise trigonométrica [64].

Portanto a proposta desta etapa consistiu em calcular a diferença entre o tamanho do pico R e do pico S em cada uma dessas derivações para verificar o direcionamento do vetor cardíaco das derivações DI e aVF . Em seguida, para encontrar o ângulo do vetor foi utilizada a relação trigonométrica

$$Ec(n) = \arctan \frac{R_{D1}(n) - |S_{D1}(n)|}{R_{aVF}(n) - |S_{aVF}(n)|},$$

onde n corresponde a cada ciclo de QRS do sinal de ECG, Ec é o eixo cardíaco calculado em graus, R_{D1} é a amplitude da onda R da derivação DI , S_{D1} é a amplitude da onda S da derivação DI , R_{aVF} é a amplitude da onda R da derivação aVF , S_{aVF} é a amplitude da onda S da derivação aVF .

Sendo assim, o eixo cardíaco foi definido como uma das características utilizadas na entrada dos sistemas classificadores explicados na Subseção 3.1.2.

3.1.2 Desenvolvimento de sistema classificador

SVM e Ensemble A proposta deste trabalho consistiu em realizar o treinamento supervisionado utilizando Máquinas de Vetores Suporte (SVM, do inglês *Support Vector Machine*) e *Ensemble* (RobustBoost), que são duas técnicas diferentes de aprendizagem supervisionada que necessitam de uma extração prévia das características. A SVM é apresentada na literatura como uma técnica utilizada na classificação dos sinais de ECG que possui bom desempenho de treinamento e resultados estatísticos significativos, em alguns casos com valores de acurácia de 99% a 100% [67, 68, 47].

Rede Neural Convolutacional Além desta proposta de extrair características previamente foi empregada a Rede Neural Convolutacional (CNN, do inglês *Convolutional Neural Network*), para a classificação dos sinais de ECG. Esse sistema é utilizado para classificação sem a necessidade de uma extração prévia das características, podendo ser utilizado na classificação de doenças cardiovasculares [69].

A literatura apresenta a relevância dos sistemas classificadores convolucionais em sinais de ECG, pois podem contribuir com a eficácia de análise de dados, favorecendo o desenvolvimento de sistemas em tempo real de monitoramento desses sinais [70]. Portanto, a proposta desta etapa consistiu em utilizar CNN na classificação dos sinais de ECG

relacionados à SM e obter as métricas de desempenho, explicadas na Subseção 3.3.2, e compará-las com os resultados obtidos pelos outros classificadores em que foram realizados extração de características anteriormente – SVM e *Ensemble*.

Portanto, para isso foram desenvolvidos algoritmos que realizavam a leitura dos sinais de ECG integrais, de 12 derivações, dos dois grupos – grupo com SM e grupo controle, treinamento e validação das CNN, seguida de análises estatísticas de desempenho.

3.2 METODOLOGIA EXPERIMENTAL

A proposta deste trabalho consistiu em utilizar sinais ECG disponibilizados nos bancos de dados apresentados na Subseção 3.2.1. Todos os algoritmos foram desenvolvidos no ambiente de desenvolvimento do Matlab, utilizando o sistema operacional Windows 10 de 64 bits, em um notebook com as seguintes especificações: placa de vídeo NVIDIA GeForce de 2 GB, Intel Core I5, 8 GB de memória RAM e 1 TB de armazenamento.

3.2.1 Banco de dados

Sinais de ECG utilizados para analisar os resultados de detecção do complexo QRS Para detectar os picos manualmente e verificar o erro obtido em relação aos resultados do algoritmo de detecção automática foi utilizado o banco de dados de teste de ECG com ruído da página do Physionet intitulada *The Noise Stress Test Database*. Para isso foi selecionado o sinal de ECG de 30 minutos de duração, da derivação V1 e de 24 SNRdB que possui frequência de amostragem de 360 Hz [71, 72].

Sinais de ECG utilizados para desenvolvimento do sistema detector de Síndrome Metabólica Neste trabalho foi realizado o estudo dos sinais de eletrocardiografia fornecidos pelo Grupo de Bioengenharia e Biofísica Aplicada – do espanhol, *Grupo de Bioingeniería y Biofísica Aplicada (GBBANet)*, que avalia as alterações nos sinais relacionadas à resistência a insulina na presença da SM [73]. Esses sinais utilizados estão apresentados em um artigo científico de um estudo publicado em 2012 na Revista Brasileira de Engenharia Biomédica [13].

Os participantes da pesquisa desse banco de dados pertencem a dois grupos distintos, sendo 15 indivíduos com SM e 10 do grupo controle constituídos por indivíduos saudáveis. Em relação a esses dados, não foi encontrada informação detalhada do diagnóstico de SM desses participantes. Em outras palavras, não foram encontradas informações antro-

ométricas individuais, tais como, idade, altura, glicose, triglicerídeo, colesterol, pressão arterial, peso. Sabe-se apenas que todos os indivíduos eram do sexo masculino, com idades de 20 a 44 anos, sem doenças cardiovasculares conhecidas, não fumavam e não faziam uso de medicamentos, e que o grupo com SM possuía obesidade, principalmente abdominal [13]. Portanto, como não foram encontradas informações relacionadas aos fatores de risco do grupo com SM, neste estudo não viável avaliar e comparar possíveis alterações no sistema cardiovascular provocadas por diferentes fatores de risco.

Esses dados foram obtidos durante o Teste Oral de Tolerância a Glicose (TOTG) realizado nos indivíduos do grupo com SM e do grupo controle. Esse procedimento consiste em um teste, realizado em jejum, em que são realizados exames de sangue antes e depois da ingestão de 75 gramas de glicose líquida. Além dos exames de sangue, foram obtidos os sinais de eletrocardiografia em 5 etapas diferentes, sendo primeiro em jejum, denominado estado basal, e os outros foram adquiridas após a ingestão de glicose com intervalos de 30 minutos, portanto foram realizadas aquisições 30, 60, 90 e 120 minutos após a ingestão da glicose [13].

Foram realizadas as aquisições dos sinais de eletrocardiografia de 12 derivações por meio do produto fornecido pela *Cardiosoft*. Cada aquisição teve duração de 15 minutos, com frequência de amostragem de 1kHz e resolução de 16 bits [13, 74].

3.2.2 Procedimentos utilizados para desenvolvimento dos sistemas classificadores

Seleção dos dados para treinamento e validação dos sistemas classificadores

Para desenvolvimento dos sistemas classificadores de SM utilizando sinais de ECG, os dados são divididos em dois grupos, um para o treinamento e outro para validação. Para isso foi realizada uma técnica estatística que avalia e compara algoritmos de aprendizagem, denominada validação cruzada. Para esta técnica optou-se pelo método k-fold, forma básica da validação cruzada apresentada na literatura como um método de estimativa precisa de desempenho [75, 76].

Nesta etapa, os dados armazenados foram inicialmente aleatorizados, em seguida, divididos em k conjuntos, sendo um deles utilizado para validação e k-1 para o treinamento. Em seguida, as etapas de treinamento e validação são realizadas em k iterações, alterando-se a cada iteração o conjunto de dados utilizado para treinar e para validar [76]. A cada iteração foram calculadas as medidas de desempenho apresentadas na Subseção 3.3.2 e ao final são calculadas as médias desses resultados.

A Figura 3.5 apresenta as etapas de desenvolvimento dos sistemas classificadores uti-

lizando k-fold em um exemplo com k igual 3, em que, os dados de entrada são inicialmente aleatorizados em seguida divididos em 3 agrupamentos para treinamento e validação.

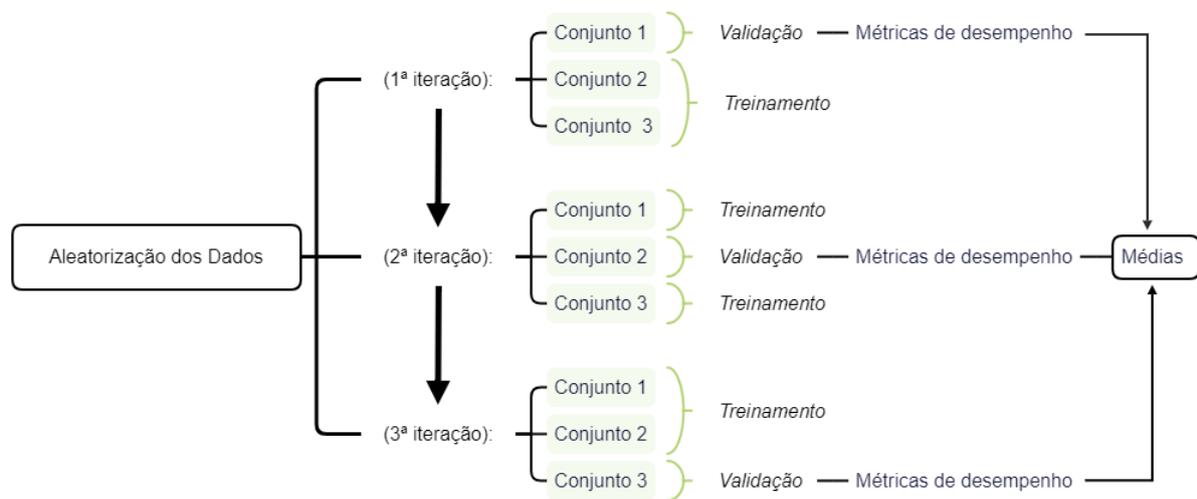


Figura 3.5. Fluxograma das etapas da metodologia experimental de treinamento e validação dos classificadores em um exemplo do k-fold com k igual a 3. Baseado em [76].

SVM e Ensemble Para o treinamento utilizando SVM, a função kernel definida foi a Função de Base Radial – do inglês, *Radial Basis Function (RBF)*, função normalmente utilizada nesse tipo de aplicação.

No caso do *Ensemble*, o tipo de classificador definido foi o *RobustBoost*, classificador binário, 100 ciclos de aprendizagem e modelos de classificação definidos como árvore; esses parâmetros foram definidos em testes preliminares, já que menos de 100 ciclos não haviam resultado em valores de acurácia de 90%.

Para a seleção dos dados desses classificadores foi utilizado o método k-fold com k igual a 10 [76].

Rede Neural Convolutacional Para o desenvolvimento da CNN foram desenvolvidos algoritmos no matlab, em que a primeira etapa consistiu em realizar a leitura dos sinais de ECG integrais de 12 derivações, dos dois grupos – grupo com SM e grupo controle, seguida do treinamento e validação.

O banco de dados fornece os sinais de apenas 25 participantes, esse número de exemplos não foi suficiente para realizar o treinamento da CNN, logo cada sinal foi dividido em 300 janelas – quantidade testada que permitiu um número suficiente de exemplos com

resultados estatisticamente significativos. Logo foram obtidos 7500 exemplos, sendo 3000 exemplos do grupo controle e 4500 do grupo com SM. Esses dados foram armazenados em uma variável com memória alocada de 4 dimensões para armazenar as 12 derivações do sinal unidimensional – sinal de ECG, sendo 300 exemplos de cada paciente. Além disso, uma outra variável recebe as classes dos sinais indicando se o sinal corresponde a um paciente do grupo com SM ou do grupo controle.

A CNN é constituída por duas etapas, etapa de extração de características e de classificação, em um total de 10 camadas. Essas camadas são mostradas na Figura 3.6.

Na etapa de extração de características, a primeira camada consiste na leitura do sinal que, em seguida, passa pela camada convolucional bidimensional com 100 filtros de tamanho 5, uma camada Unidade Linear Retificada (ReLU, do inglês *Rectified Linear Unit*), seguida de uma camada *maxpooling* bidimensional de passo 1. Em seguida, a rede é constituída por uma segunda etapa convolucional bidimensional com 100 filtros de tamanho 8, novamente por uma camada ReLU, seguida por uma camada *maxpooling* bidimensional de passo 2, estes parâmetros foram estabelecidos, pois geravam resultados com valores significativos, de exatidão entre 90% e 100%.

As camadas de classificação são constituídas por uma camada densa para classificação bidimensional, uma camada *softmax* e uma camada de saída de classificação para uma rede neural.

Para o treinamento da CNN foram definidas as seguintes opções que obtiveram os melhores resultados estatísticos: descida gradiente estocástica com impulso; taxa de aprendizagem inicial: 0,0001; taxa de aprendizagem inicial multiplicada por 0,1 a cada 8 épocas; fator de regularização L2 de 0,0004 (reduz o *overfittting*); número máximo de épocas: 5; tamanho do mini-lote usado para cada iteração: 30 (avalia o gradiente da função de perda e atualiza os pesos).

3.3 ANÁLISES ESTATÍSTICAS

As análises estatísticas realizadas nesta proposta são divididas em dois subgrupos, o primeiro consiste em análises sobre as características extraídas dos sinais de ECG e o segundo consiste em análises sobre os resultados de classificação dos diferentes classificadores utilizados nesta proposta, SVM, RobustBoost e CNN. É possível visualizar na Figura 3.7 as etapas das análises estatísticas realizadas.

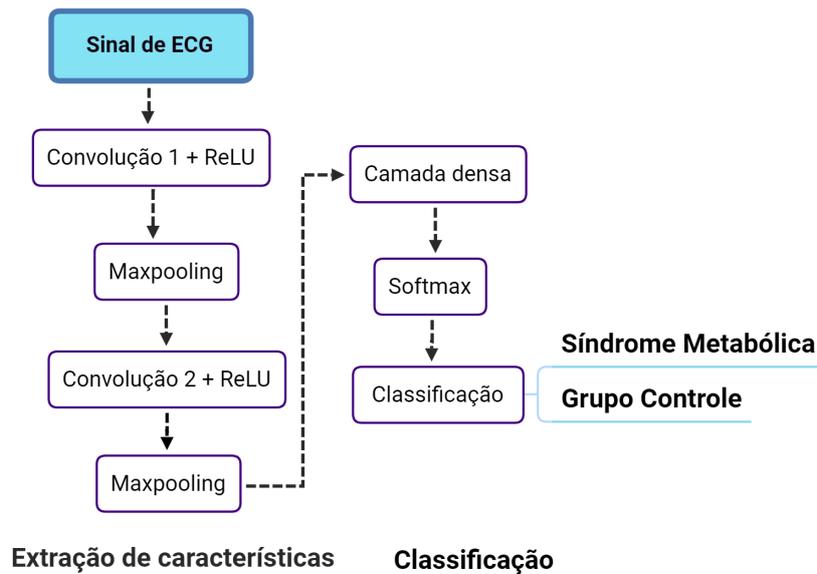


Figura 3.6. Camadas da Rede Neural Convolucional projetada para o treinamento e classificação dos sinais de ECG em dois grupos (grupo controle e grupo com Síndrome Metabólica). Fonte própria.

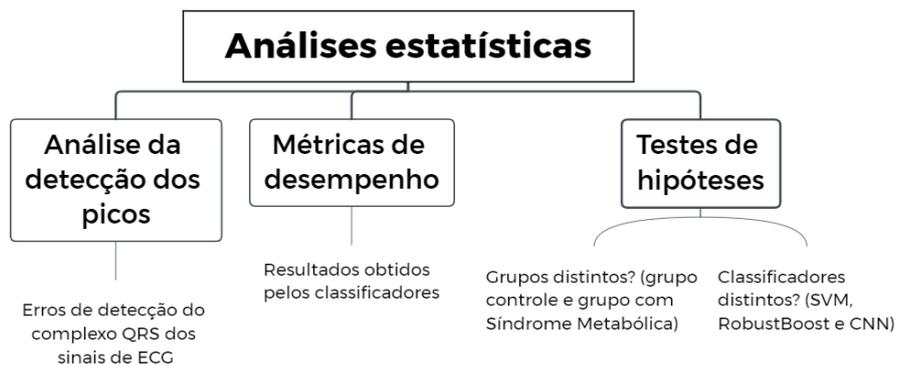


Figura 3.7. Etapas das análises estatísticas realizadas nas características extraídas dos sinais de ECG e nos resultados obtidos pelos sistemas classificadores. Fonte própria.

3.3.1 Análises estatísticas dos resultados obtidos pelos algoritmos de detecção dos picos do sinal de ECG

O cálculo do erro de detecção consiste em

$$E(n) = |Lpd(n) - Lpr(n)|,$$

sendo E o erro detectado, em valor absoluto, Lpd a localização do pico (Q, R ou S) detectada pelo algoritmo, Lpr a localização do pico considerada real (extraída manualmente ou fornecida pelo banco de dados) e n a quantidade de picos detectados conforme o tamanho do sinal. Logo em seguida foi calculada a média de $E(n)$.

Para verificar a eficácia dos algoritmos de detecção dos picos foi utilizada uma técnica de análise gráfica desenvolvida por Bland e Altman em 1986, que utiliza medidas de estatística descritiva, tais como a média e o desvio padrão. Esse tipo de gráfico foi criado para avaliar e comparar resultados de diferentes técnicas aplicadas na área da saúde [63].

A proposta do gráfico de Bland-Altman é apresentar a relação entre as diferenças e a média entre os resultados das duas técnicas. A média é utilizada como uma medida de aproximação do valor real. Portanto o gráfico de Bland-Altman consiste em representar graficamente o valor da média em função da diferença, além da média das diferenças e os limites superior e inferior definidos conforme a distribuição dos dados [63].

Para testar a normalidade dos dados, foi utilizado o teste de Kolmogorov-Smirnov – explicado em detalhes na Subseção 3.3.3. Segundo esse teste, os dados não seguiam distribuição gaussiana, portanto foram utilizadas estatísticas não paramétricas, logo os limites superior e inferior foram calculados como

$$li = 1,45IQR$$

e

$$ls = 1,45IQR,$$

respectivamente, onde li é o limite inferior, ls é o limite superior e IQR é o intervalo interquartil – do inglês *Interquartile Range*, medida estatística que avalia a dispersão dos dados.

A Figura 3.8 mostra um exemplo do gráfico de Bland-Altman de um estudo que comparava dois métodos diferentes de medição da taxa de pico de fluxo expiratório. Logo o intuito desse gráfico é verificar quanto que o erro varia comparando duas técnicas diferentes. Neste caso, de distribuição gaussiana, os limites inferiores e superiores correspondem a

$$li = d - 2s,$$

$$ls = d + 2s,$$

respectivamente, onde d é a média e s é o desvio padrão das diferenças.

3.3.2 Métricas de desempenho dos sistemas classificadores testados

Após a obtenção dos resultados de validação dos sistemas classificadores, referentes aos diferentes estágios de aquisição dos sinais de ECG do banco de dados apresentado na Subseção 3.2.1, para as análises estatísticas normalmente utilizadas neste tipo de estudo

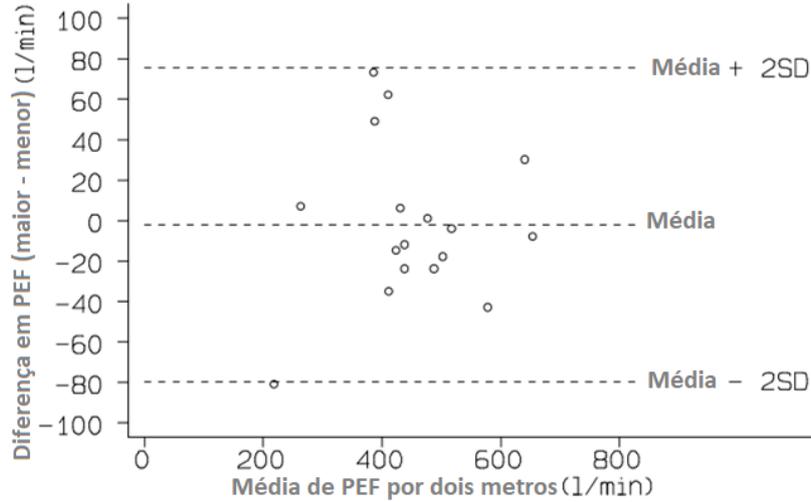


Figura 3.8. Exemplo do gráfico de Bland Altman, onde foi medida a taxa de pico de fluxo expiratório (PEF). Adaptado de [63].

foi utilizada a matriz de confusão da Figura 3.9.

Classes reais	Classes geradas pelo classificador	
	Indivíduos com Síndrome Metabólica (SM)	Indivíduos saudáveis, do grupo controle (GC)
Indivíduos com Síndrome Metabólica (SM)	VP	FN
Indivíduos saudáveis, do grupo controle (GC)	FP	VN

Figura 3.9. Matriz de confusão desenvolvida para análise das métricas de desempenho dos sistemas classificadores nos sinais de ECG do banco de dados conforme Subseção 3.2.1. Sendo VP quantidade de verdadeiros positivos, FN quantidade de falsos negativos, FP quantidade de falsos positivos e VN quantidade de verdadeiros negativos. Fonte própria.

Em que VP é a quantidade de verdadeiro positivo, classificações correta de sinais de indivíduos com SM; VN é a quantidade de verdadeiros negativos, classificações correta de sinais de indivíduos do grupo controle; FP é a quantidade de falso positivo, sinais classificados erroneamente com SM; e FN é quantidade de falsos negativos, sinais classificados erroneamente como grupo controle. Portanto, os valores obtidos nas posições da diagonal correspondem às classificações corretas.

A partir da matriz de confusão, as seguintes métricas de desempenho foram determinadas: taxa de falsos positivos, taxa de falsos negativos, valor preditivo positivo, valor

preditivo negativo, sensibilidade, especificidade, exatidão e medidas-F, representadas por

$$TFP = \frac{FP}{FP + VN},$$

$$TFN = \frac{FN}{VP + FN},$$

$$VPP = \frac{VP}{VP + FP},$$

$$VPN = \frac{VN}{VN + FN},$$

$$S = \frac{VP}{VP + FN},$$

$$Es = \frac{VN}{VN + FP},$$

$$Ex = \frac{VP + VN}{VP + VN + FP + FN},$$

$$MF = \frac{2(VPP \cdot S)}{VPP + S},$$

respectivamente, onde TFP é a taxa de falsos positivos (erro tipo I), TFN é a taxa de falsos negativos (erro tipo II), VPP é o valor preditivo positivo, VPN é o valor preditivo negativo, S é a sensibilidade, Es é a especificidade, Ex é a exatidão e MF é a medida-F.

3.3.3 Testes estatísticos de hipótese

Este trabalho consistiu na identificação de classes relacionadas a um possível diagnóstico de SM utilizando diferentes técnicas de classificação explicadas na Seção 3.1. Sendo assim, a análise estatística deste estudo consistiu na inferência estatística, em que uma amostra de valores, características extraídas dos sinais de ECG, representa a população. Portanto, foram utilizados testes para avaliar a hipótese de que é possível diferenciar indivíduos com SM a partir dos sinais de ECG. Logo, a hipótese nula consiste em não ser possível diferenciar os indivíduos do grupo com SM. A ideia foi verificar a rejeição da hipótese nula, com nível de confiança de 5%. Esses testes estatísticos foram realizados no matlab.

Primeiramente, para verificar se os dados diferem de uma distribuição normal foi aplicado o teste de normalidade de Kolmogorov-Smirnov. Para este teste, as hipóteses testadas são

H_0 – Hipótese nula: os dados seguem distribuição gaussiana; e

H1 – Hipótese alternativa: os dados não seguem distribuição gaussiana.

Esse estudo consiste em um caso de amostras independentes, pois os dados extraídos dos sinais de ECG de cada grupo são distintos, ou seja, os dados de um grupo não interferem no outro. Portanto, como as características não seguiam distribuição gaussiana, foi utilizado o teste estatístico não paramétrico para amostras independentes de Mann-Whitney, em que foram testadas as seguintes hipóteses

H0 – Hipótese nula: as características extraídas dos sinais de ECG pertencem a indivíduos de grupos similares; e

H1 – Hipótese alternativa: as características extraídas dos sinais de ECG pertencem a indivíduos de grupos distintos.

Além disso, esses testes estatísticos foram aplicados nos resultados obtidos pelos diferentes classificadores a fim de verificar se as diferenças obtidas são estatisticamente significativas, neste caso as hipóteses a serem testadas foram

H0 – Hipótese nula: não existem diferenças estatisticamente significativas entre os resultados obtidos pelos diferentes classificadores;

H1 – Hipótese alternativa: existem diferenças estatisticamente significativas entre os resultados obtidos pelos diferentes classificadores.

4 RESULTADOS E DISCUSSÃO

4.1 ANÁLISES ESTATÍSTICAS DOS RESULTADOS DE DETECÇÃO DOS PICOS DOS SINAIS DE ECG

A fim de validar os resultados de detecção dos picos das ondas Q, R e S e verificar possíveis erros entre os valores dos picos detectados e os valores reais, conforme explicado na Subseção 3.1.1, foi gerado no matlab o gráfico de Bland-Altman para comparar os resultados obtidos pelos algoritmos desenvolvidos – detecção automática e manual [77]. Neste algoritmo foi realizado teste de normalidade das amostras utilizando o teste de Kolmogorov-Smirnov. Verificou-se que as distribuições dos dados, localizações das ondas Q, R e S, não seguem distribuição normal, portanto foi utilizada análise estatística para dados não paramétricos. As Figuras 4.1, 4.2 e 4.3 apresentam os gráficos de Bland-Altman referentes as detecções dos picos Q, R e S, respectivamente. Observe que os pontos distantes de zero nestes gráficos correspondem a erros de detecção que ocorreram nas regiões de transição desses valores, provocando um aumento nas diferenças entre os resultados obtidos pelos dois algoritmos.

Logo, quando o algoritmo de detecção manual não realiza a detecção de um pico, ocorre aumento nos resultados das diferenças entre os picos obtidos pelos dois algoritmos a partir dessa região, visto que a subtração entre os resultados do algoritmo de detecção automática torna-se, a partir desse erro, relacionada aos resultados obtidos pelos picos do ciclo do complexo QRS anterior pelo algoritmo de detecção manual. Além disso, de forma equivalente, uma detecção a mais provoca redução nos valores das diferenças seguintes obtidas.

Diante disso, analisando os gráficos das Figuras 4.1 a 4.3, foi possível observar as regiões onde ocorreram erros nas detecções dos picos. Nas Figuras 4.1 e 4.3, os erros nas regiões próximas de 300 e de 1500 segundos foram provocados devido ao ruído presente nestas regiões que ocasionaram aumento das diferenças entre as posições dos picos detectados. Esses ruídos são relacionados a artefatos de movimento que foram adicionados aos sinais, que são fornecidos pelo banco de dados de teste de ECG com ruído apresentado

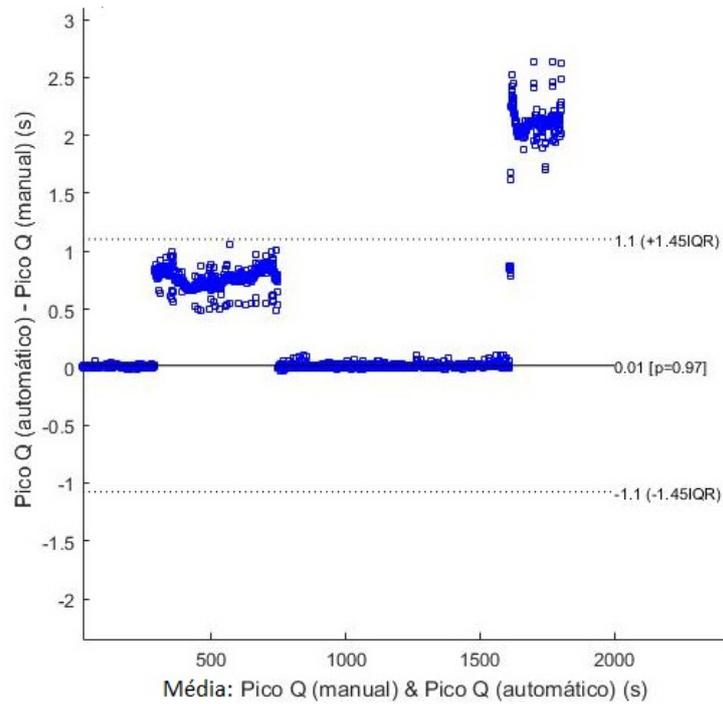


Figura 4.1. Média em função da diferença entre as localizações dos picos Q utilizando o algoritmo de detecção automático e o algoritmo de detecção manual. Fonte própria.

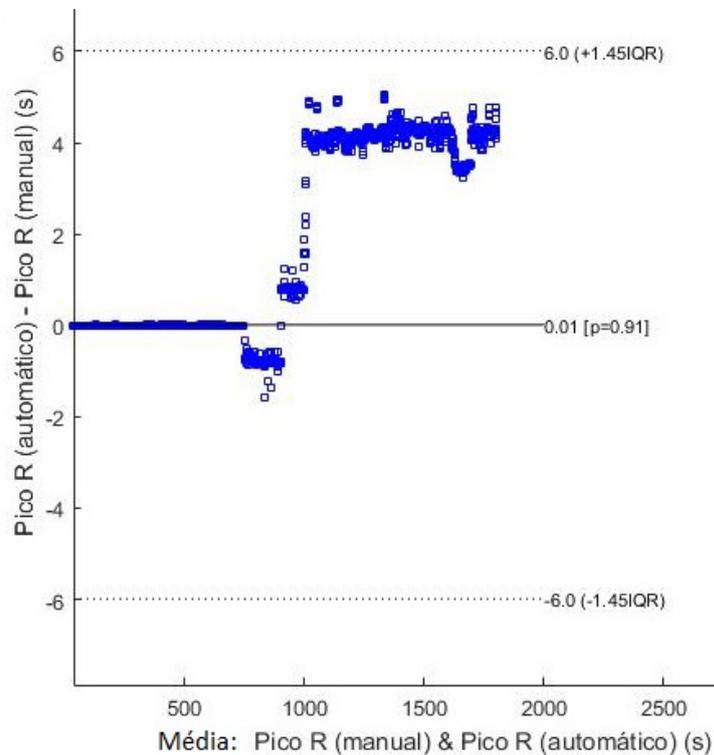


Figura 4.2. Média em função da diferença entre as localizações dos picos R utilizando o algoritmo de detecção automático e o algoritmo de detecção manual. Fonte própria.

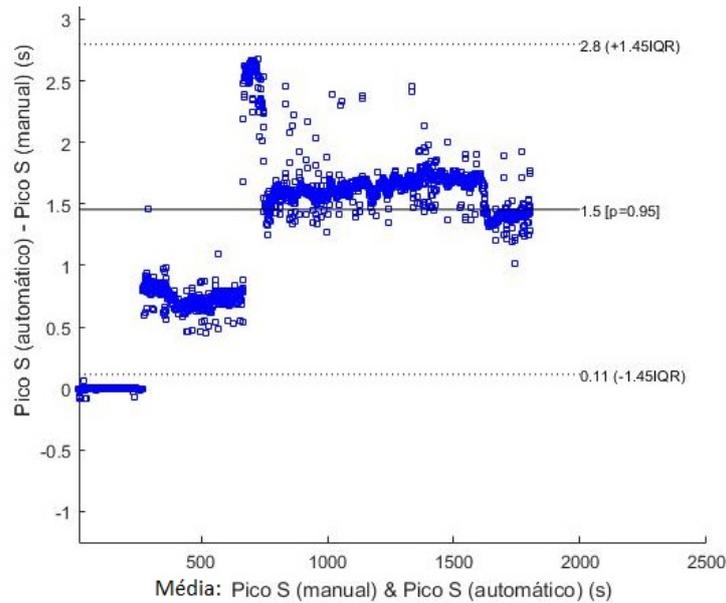


Figura 4.3. Média em função da diferença entre as localizações dos picos S utilizando o algoritmo de detecção automático e o algoritmo de detecção manual. Fonte própria.

na Subseção 3.2.1. Além disso, na região do sinal próxima a 720 segundos foram encontradas alterações nos gráficos das Figuras 4.1 a 4.3, pois foi detectado equivocadamente um complexo QRS em uma região de ruído pelo método de detecção automática.

O gráfico de Bland Altman é uma estratégia que analisa se diferenças obtidas por duas técnicas distintas são significativas e se isso pode interferir em decisões aplicadas em diagnóstico [63]. Logo, conforme os resultados obtidos, utilizando essa metodologia para detecção dos complexos QRS, o diagnóstico pode ser comprometido na presença de ruídos, gerados por artefatos de movimento, pois nessa situação o algoritmo possui seu desempenho prejudicado.

4.2 MÉTRICAS ESTATÍSTICAS DE DESEMPENHO DOS SISTEMAS CLASSIFICADORES

Após a extração das características dos sinais de ECG, elas foram utilizadas no treinamento e validação de sistemas classificadores. Na primeira etapa foram utilizados Máquinas de Vetores Suporte (SVM, do inglês *Support Vector Machine*) e o *Ensemble* do tipo *RobustBoost*.

Os sinais de ECG do banco de dados utilizado, conforme explicado na Subseção 3.2.1, foram adquiridos em cinco estágios diferentes, primeiramente em estágio basal onde os participantes estavam em jejum e depois foram realizadas quatro etapas em intervalos de

30 minutos após a ingestão da glicose líquida.

Para analisar os resultados estatísticos dos sistemas classificadores foram utilizadas as seguintes métricas de desempenho, taxa de falsos positivos (*TFP*), taxa de falsos negativos (*TFN*), valor preditivo positivo (*VPP*), valor preditivo negativo (*VPN*), sensibilidade (*S*), exatidão (*Ex*), especificidade (*Es*) e medidas-F (*MF*).

A Tabela 4.1 apresenta os resultados obtidos utilizando características extraídas de apenas uma derivação do sinal de ECG (derivação *DI*), enquanto que a Tabela 4.2 apresenta os resultados obtidos avaliando duas características dos sinais de ECG (derivação *DI* e *aVF*, incluindo o eixo cardíaco). Por meio desses resultados não foram observadas diferenças significativas, portanto, segundo essa metodologia adotada, o aumento das características não contribuiu com a melhora dos resultados obtidos.

Avaliando a Tabela 4.2, é possível afirmar que os resultados de detecção da SM possuíram valores estatisticamente significativos, tais como, *TFP* abaixo de 11,2%, *TFN* abaixo de 11,5%, *VPP* de até 99,7%, *VPN* de até 99,9%, *S* de até 96,1%, *Ex* de até 97%, *Es* de até 99,9% e *MF* de até 97,2%.

Observou-se também que quanto maior a porcentagem de falsos positivos, ou seja, o erro tipo I, menor foi a porcentagem de falsos negativos – erro tipo II, menor foi o valor preditivo positivo e maior foi o valor preditivo negativo. E os resultados das medidas-F acima de 91,5% indicam que a especificidade e a exatidão possuem valores estatisticamente significativos, acima de 81% e 88,5%, respectivamente.

Além disso, com o intuito de comparar os resultados da SVM e do *RobustBoost* foi gerado o histograma, mostrado na Figura 4.4, das exatidões obtidas no desenvolvimento dos classificadores com os sinais de ECG pertencentes ao estágio basal. Nesta figura, é possível notar que SVM obteve resultados superiores ao *RobustBoost*, com o valor de exatidão de maior frequência de aproximadamente 94%, enquanto que o *RobustBoost* obteve valores de exatidão próximos de 89%.

Com a finalidade de comparar esses classificadores com a CNN, foi realizado um teste de treinamento e validação, utilizando o k-fold com k igual a 10, para cada método de aprendizagem de máquina – SVM, *RobustBoost* e CNN. Nesta etapa foram utilizados os sinais de ECG do estágio basal fornecidos pelo banco de dados. Além disso, com o intuito de comparar resultados obtidos avaliando diferentes quantidades de derivações dos sinais de ECG, no teste realizado com SVM e no *RobustBoost* foram utilizadas as derivações *DI* e *aVF*, enquanto que no teste realizado com a CNN foram utilizadas 12 derivações dos sinais de ECG. A Tabela 4.3 apresenta as médias das medidas estatísticas obtidas.

Tabela 4.1. Média e desvio padrão das métricas de desempenho em porcentagem dos sistemas classificadores analisados em um teste com dados selecionados aleatoriamente, utilizando k-fold com k igual a 10, em apenas uma derivação dos sinais de ECG (derivação *DI*). Sendo *TFP* a taxa de falsos positivos (erro tipo I), *TFN* é a taxa de falsos negativos (erro tipo II), *VPP* é o valor preditivo positivo, *VPN* é o valor preditivo negativo, *S* é a sensibilidade, *Ex* é a exatidão, *Es* é a especificidade e *MF* é a medida-F.

Estágio:	basal	30 minutos	60 minutos	90 minutos	120 minutos
SVM :					
<i>TFP</i> (%)	6,4 ± 1,4	5,8 ± 4,0	9,6 ± 4,9	11,8 ± 4,8	5,7 ± 4,4
<i>TFN</i> (%)	0,2 ± 0,7	5,1 ± 3,5	2,3 ± 2,3	5,7 ± 4,8	4,6 ± 4,5
<i>VPP</i> (%)	93,6 ± 1,4	94,2 ± 4,0	90,4 ± 4,9	88,2 ± 4,8	94,3 ± 4,4
<i>VPN</i> (%)	99,3 ± 2,1	92,5 ± 6,3	96,1 ± 3,8	94,1 ± 4,8	93,0 ± 6,6
<i>S</i> (%)	99,8 ± 0,7	94,9 ± 3,5	97,7 ± 2,3	94,2 ± 4,8	95,4 ± 4,5
<i>Ex</i> (%)	95,7 ± 1,2	93,6 ± 3,2	92,4 ± 3,6	90,6 ± 3,7	93,6 ± 2,7
<i>Es</i> (%)	88,8 ± 3,4	91,7 ± 4,6	84,6 ± 7,0	86,6 ± 6,9	91,1 ± 7,1
<i>MF</i> (%)	96,6 ± 0,8	94,5 ± 3,0	93,8 ± 3,2	91,0 ± 3,5	94,7 ± 2,4
RobustBoost:					
<i>TFP</i> (%)	11,2 ± 4,4	7,4 ± 5,2	6,4 ± 6,8	13,4 ± 12,1	11,7 ± 6,1
<i>TFN</i> (%)	5,2 ± 4,6	8,8 ± 4,4	5,6 ± 4,3	4,5 ± 3,7	6,0 ± 3,7
<i>VPP</i> (%)	88,8 ± 4,3	92,6 ± 5,2	93,6 ± 6,8	86,6 ± 12,0	88,2 ± 6,1
<i>VPN</i> (%)	89,7 ± 10,7	86,3 ± 7,2	89,3 ± 8,5	94,9 ± 4,4	90,1 ± 6,3
<i>S</i> (%)	94,8 ± 4,6	91,2 ± 4,4	94,3 ± 4,3	95,5 ± 3,7	94,0 ± 3,7
<i>Ex</i> (%)	89,3 ± 4,6	90,0 ± 3,7	92,4 ± 4,6	90,2 ± 8,1	89,1 ± 4,5
<i>Es</i> (%)	80,5 ± 8,0	88,7 ± 7,7	89,7 ± 8,7	85,4 ± 12,9	82,3 ± 8,6
<i>MF</i> (%)	91,6 ± 3,5	91,8 ± 3,0	93,8 ± 4,3	90,5 ± 7,9	90,9 ± 4,0

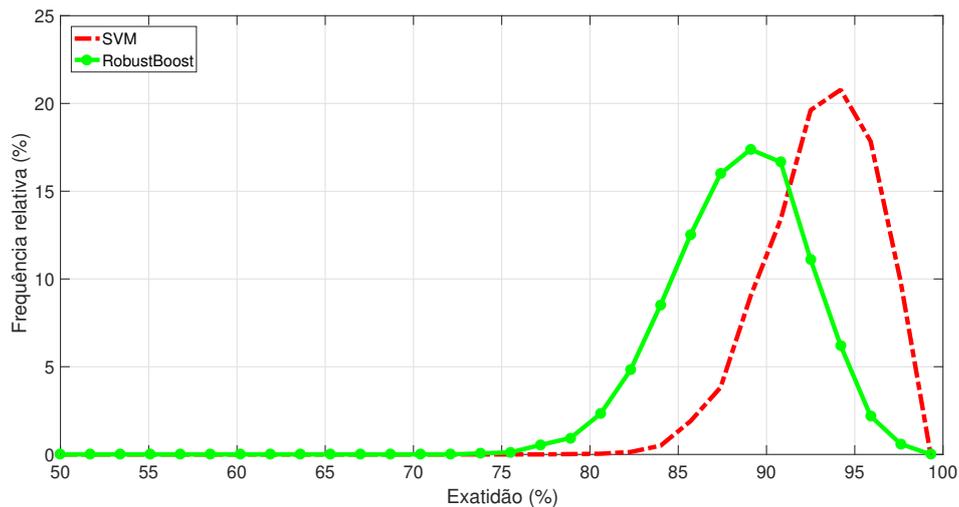


Figura 4.4. Exatidões em função das frequências relativas obtidas pela SVM e pelo *RobustBoost* na classificação dos sinais de ECG adquiridos durante o estágio Basal, utilizando k-fold com k igual a 10 em 500 testes. Fonte própria.

Tabela 4.2. Média e desvio padrão das métricas de desempenho em porcentagem dos sistemas classificadores analisados em um teste com 500 iterações de treinamentos e validações com dados selecionados aleatoriamente, utilizando k-fold com k igual a 10, em duas derivações dos sinais de ECG (derivação *DI* e *aVF*). Sendo *TFP* a taxa de falsos positivos (erro tipo I), *TFN* é a taxa de falsos negativos (erro tipo II), *VPP* é o valor preditivo positivo, *VPN* é o valor preditivo negativo, *S* é a sensibilidade, *Ex* é a exatidão, *Es* é a especificidade e *MF* é a medida-F.

Estágio:	basal	30 minutos	60 minutos	90 minutos	120 minutos
SVM :					
<i>TFP</i> (%)	1,5 ± 1,9	0,2 ± 0,8	0,3 ± 0,9	0,1 ± 0,8	0,3 ± 0,9
<i>TFN</i> (%)	7,1 ± 4,1	8,2 ± 4,7	7,0 ± 4,0	11,5 ± 6,1	6,8 ± 4,1
<i>VPP</i> (%)	98,5 ± 1,9	99,7 ± 0,8	99,7 ± 0,9	84,7 ± 8,0	99,7 ± 0,9
<i>VPN</i> (%)	85,8 ± 8,0	86,5 ± 7,7	86,7 ± 7,3	99,9 ± 0,8	87,7 ± 7,1
<i>S</i> (%)	92,9 ± 4,1	91,8 ± 4,7	93,0 ± 4,0	84,7 ± 8,0	93,1 ± 4,1
<i>Ex</i> (%)	94,1 ± 3,1	94,5 ± 3,2	95,0 ± 2,3	88,5 ± 6,1	95,2 ± 2,8
<i>Es</i> (%)	96,7 ± 4,1	99,6 ± 1,5	99,5 ± 1,8	99,9 ± 1,1	99,4 ± 1,7
<i>MF</i> (%)	95,5 ± 2,4	95,5 ± 2,6	96,2 ± 2,2	93,7 ± 3,5	96,2 ± 2,2
<i>RobustBoost:</i>					
<i>TFP</i> (%)	11,2 ± 4,7	3,1 ± 3,5	4,8 ± 4,1	1,5 ± 2,9	6,5 ± 8,1
<i>TFN</i> (%)	5,4 ± 3,6	5,5 ± 4,9	8,1 ± 5,7	3,8 ± 4,8	4,8 ± 5,0
<i>VPP</i> (%)	88,8 ± 4,7	96,9 ± 3,5	95,2 ± 4,1	98,4 ± 2,9	93,5 ± 8,1
<i>VPN</i> (%)	90,4 ± 6,3	91,3 ± 7,8	84,7 ± 11,3	95,3 ± 5,8	92,4 ± 7,9
<i>S</i> (%)	94,6 ± 3,6	94,5 ± 4,9	91,9 ± 5,7	96,1 ± 4,8	95,2 ± 5,0
<i>Ex</i> (%)	89,3 ± 3,8	94,7 ± 4,2	91,4 ± 4,5	97,0 ± 3,3	93,0 ± 6,9
<i>Es</i> (%)	81,0 ± 7,7	95,1 ± 5,6	91,1 ± 7,2	98,1 ± 3,6	90,2 ± 11,9
<i>MF</i> (%)	91,5 ± 3,1	95,6 ± 3,5	93,4 ± 3,5	97,2 ± 3,1	94,2 ± 6,0

Tabela 4.3. Média e desvio padrão das métricas de desempenho em porcentagem dos sistemas classificadores analisados em um teste, utilizando o método k-fold, com k = 10, utilizando SVM, *RobustBoost* e CNN. Sendo *TFP* a taxa de falsos positivos (erro tipo I), *TFN* é a taxa de falsos negativos (erro tipo II), *VPP* é o valor preditivo positivo, *VPN* é o valor preditivo negativo, *S* é a sensibilidade, *Ex* é a exatidão, *Es* é a especificidade e *MF* é a medida-F.

	SVM	<i>RobustBoost</i>	CNN
<i>TFP</i> (%)	2,0% ± 2,9	11,4% ± 4,3	0,7% ± 0,6
<i>TFN</i> (%)	7,0% ± 4,8	5,5% ± 4,0	0,4% ± 0,6
<i>VPP</i> (%)	98,0% ± 2,9	88,6% ± 4,3	99,3% ± 0,6
<i>VPN</i> (%)	86,3% ± 9,4	90,4% ± 6,9	99,3% ± 1,0
<i>S</i> (%)	93,0% ± 4,8	94,5% ± 4,0	99,6% ± 0,6
<i>Ex</i> (%)	93,8% ± 4,3	89,3% ± 3,5	99,3% ± 0,4
<i>Es</i> (%)	95,8% ± 5,7	80,9% ± 6,0	98,9% ± 1,0
<i>MF</i> (%)	95,4% ± 3,3	91,4% ± 3,4	99,4% ± 0,3

Esses resultados apresentados na Tabela 4.3 indicam que foi possível obter métricas de desempenho com resultados estatisticamente significativos utilizando esses métodos de aprendizagem avaliando diferentes quantidades e tipos de derivações dos sinais de ECG, com exatidão acima de 89%, para detecção de Síndrome Metabólica. Sendo que, conforme Tabela 4.4, utilizando CNN os resultados foram superiores.

Tabela 4.4. Moda aproximada da exatidão obtida pelos diferentes classificadores em 10 testes realizados utilizando k-fold com k igual a 3 nos sinais de ECG adquiridos durante o estágio Basal.

Moda da exatidão	
SVM (%)	94%
RobustBoost (%)	89%
CNN (%)	98%

Entretanto, conforme mostrado na Tabela 4.5, a CNN exigiu maior custo computacional durante a etapa de treinamento do sistema classificador, em comparação com o *RobustBoost* e a SVM.

Tabela 4.5. Tempo aproximado de execução dos algoritmos, utilizando a técnica k-fold para separar a quantidade de dados para treinamento e validação, com k igual a 10, em um teste realizado com cada sistema classificador, utilizando o mesmo *hardware* e mesmo sistema operacional.

Classificador	tempo (treinamento)	tempo (validação)
SVM :	0,3 segundos	0,02 segundos
RobustBoost:	11 segundos	0,06 segundos
CNN :	23 minutos	10 segundos

Diante desses resultados apresentados nesta Seção, e visto que o conjunto de fatores da Síndrome Metabólica (SM) ocasiona alterações cardiovasculares, observou-se neste estudo que é possível detectar a SM com resultados estatisticamente significativos, com exatidão acima de 90%, por meio de sinais de Eletrocardiografia, utilizando diferentes abordagens. Incluindo sistemas que necessitam de uma etapa prévia de extração de características desses sinais utilizadas no desenvolvimento de sistemas classificadores, tais como *RobustBoost* e SVM, e também utilizando Redes Neurais Convolucionais que realizam esse procedimento automaticamente. Essas abordagens de detecção consistem em análises de apenas uma derivação – derivação *D1*, duas derivações – derivações *D1* e *aVF*, e também incluindo as 12 derivações relacionadas a eletrocardiógrafos convencionais utilizados em clínicas.

Apesar desses resultantes, este estudo possui limitações relacionadas a caracterização

da SM, pois não foram encontradas descrições antropométricas dos participante da pesquisa correspondentes ao sinais de ECG fornecidos pelo banco de dados, tais como, idade, altura, glicose, triglicérideo, colesterol, pressão arterial, peso. Portanto, neste estudo, não foi possível analisar quais fatores de risco da SM e a influência que cada um exerce em alterações no funcionamento do sistema cardiovascular.

4.2.1 Análise dos resultados obtidos pelos classificadores

Os resultados obtidos exibem desempenho dos sistemas classificadores, com valores estatisticamente significativos. Dentre esses resultados, incluem-se, por exemplo, exatidão de até 99,3%, indicando que o sistema pode ser utilizado para detecção de indivíduos com SM e de indivíduos saudáveis utilizando apenas os sinais de ECG. Além disso, a sensibilidade de até 99,6% aponta que é difícil não ser detectada a SM em indivíduos que possuem a síndrome. O valor preditivo positivo acima de 88,6% indica a taxa de acerto da detecção de indivíduos com SM. E o valor preditivo negativo acima de 86,3% demonstra a taxa de acerto dos indivíduos que foram detectados como saudáveis.

É importante mencionar que esses resultados indicam que os sistemas classificadores diferenciam sinais de ECG de indivíduos com SM de indivíduos saudáveis. Logo, os desvios provocados nos sinais de ECG em indivíduos com SM percebidos pelos sistemas são alterações não específicas desta síndrome.

Sendo assim, esses resultados estatísticos obtidos pelos sistemas classificadores sugerem que poderiam ser utilizados sinais de ECG em um sistema para detectar a presença da SM para indicar pacientes que têm que ser encaminhados para o médico para realização de exames clínicos em uma análise mais detalhada dos fatores de risco da SM. Isso porque a evolução tecnológica na área médica, com a inclusão desses sistemas automáticos, não substitui os profissionais da saúde. Esses sistemas atuam como ferramentas de auxílio, sendo os médicos responsáveis pelo diagnóstico final [78].

Comparando esses diferentes sistemas classificadores, observa-se que a CNN obteve resultados estatísticos superiores. Uma das vantagens dessa técnica, que tem sido apresentada no Estado da Arte, é a extração de características ser realizada automaticamente [35]. Ou seja, não existe a necessidade de utilizar técnicas de processamento de sinais previamente. Além disso, nas demais técnicas são utilizadas apenas algumas informações selecionadas dos sinais de ECG, relacionadas ao complexo QRS. Logo, a hipótese é que não foram incluídas outras informações que poderiam ser significantes para a detecção da SM. Como, por exemplo, intervalos QT e QTc, que são apontados em estudos como características que se alteram na presença da diabetes [60]. Além da onda P e da onda

T, que podem ser alteradas na presença da SM [6].

Já a CNN utiliza o sinal de ECG completo e extrai, automaticamente, características que são significativas para diferenciar as classes [19]. Logo, a hipótese é que a CNN obteve resultados superiores, pois características relevantes para a detecção da SM não foram desconsideradas. Portanto, conclui-se que a CNN, em comparação com as demais técnicas utilizadas neste trabalho, é a estratégia mais indicada para ser utilizada como uma ferramenta de detecção de SM por meio de sinais de ECG. Apesar da etapa de treinamento da CNN exigir maior custo computacional, para um diagnóstico em tempo real, um sistema classificador de SM já estaria previamente treinado, de forma que este custo não interferiria.

4.3 TESTES ESTATÍSTICOS DE HIPÓTESES

Segundo o teste de Kolmogorov-Smirnov realizado nas características extraídas dos sinais de ECG, os dados não seguiam distribuição gaussiana, $p < 0,01$. Portanto foram utilizadas estatísticas não paramétricas para os testes de hipóteses. A Tabela 4.6 apresenta os valores p para os testes de hipóteses realizados nas características extraídas dos sinais de ECG mostradas na Tabela C.1 no Apêndice C.

Segundo a literatura alterações cardiovasculares podem ser ocasionadas em indivíduos com SM, logo parâmetros extraídos dos sinais de ECG, tais como, informações relacionadas aos batimentos cardíacos, às ondas P, ao complexo QRS e à onda T obtidas por análises dos sinais de ECG podem ser associados à presença da SM [6, 12, 13]. Além disso, indivíduos com SM podem ter hipertrofia ventricular esquerda e comprometimento na função diastólica, essas alterações poderiam indicar um desvio do eixo em indivíduos com SM [65, 66]. Logo, os resultados obtidos neste trabalho concordam com informações apresentadas na literatura.

Os testes de hipóteses apresentados na Tabela 4.6 indicam que as características extraídas das derivações DI e aVF dos sinais de ECG, a partir da detecção do complexo QRS, relacionadas às médias e às variâncias das relações e intervalos entre os picos Q, R e S consecutivos, além do eixo cardíaco, permitem rejeitar a hipótese nula – $p < 0,01$. Em outras palavras, é possível utilizar essas características para diferenciar indivíduos do grupo controle com indivíduos do grupo com Síndrome Metabólica.

Além destes testes, com intuito de comparar os três classificadores – SVM, *Robust-Boost* e CNN – foram realizados testes de hipóteses com os valores de exatidão obtidos. Estes testes indicaram que os dados seguem distribuições distintas. Portanto existem

Tabela 4.6. Valor p do resultado do teste de hipótese de cada característica extraída dos sinais de ECG, relacionados ao grupo controle e ao grupo com SM, dos diferentes estágios do TOTG.

	basal	30 min	60 min	90 min	120 min
1	0,0	0,0	0,0	0,47	0,0
2	0,0	0,0	0,0	0,25	0,0
3	0,0	0,0	0,0	0,17	0,0
4	0,0	0,0	0,0	0,33	0,0
5	0,0	0,0	0,0	0,0	0,0
6	0,0	0,0	0,0	0,0	0,0
7	0,0	0,0	0,0	0,0	0,0
8	0,0	0,0	0,0	0,0	0,0
9	0,0	0,0	0,0	0,0	0,0
10	0,0	0,0	0,0	0,0	0,0
11	0,0	0,0	0,0	0,0	0,0
12	0,0	0,0	0,0	0,0	0,0
13	0,20	0,0	0,0	0,0	0,0
14	0,0	0,0	0,0	0,0	0,0
15	0,29	0,47	0,23	0,0	0,0
16	0,10	0,0	0,0	0,0	0,0
17	0,0	0,0	0,0	0,0	0,0
18	0,0	0,0	0,0	0,0	0,0
19	0,0	0,0	0,0	0,0	0,0
20	0,0	0,0	0,0	0,0	0,0
21	0,0	0,0	0,0	0,0	0,0
22	0,0	0,0	0,0	0,0	0,0
23	0,0	0,0	0,0	0,0	0,0
24	0,0	0,0	0,48	0,10	0,04
25	0,0	0,0	0,42	0,11	0,04
26	0,0	0,0	0,60	0,07	0,10
27	0,0	0,0	0,66	0,06	0,10
28	0,0	0,0	0,0	0,0	0,0
29	0,01	0,0	0,0	0,0	0,0
30	0,70	0,12	0,17	0,0	0,02

diferenças estatisticamente significativas, $p < 0,01$, entre os resultados de exatidões dos diferentes classificadores.

5 CONCLUSÃO

O conjunto de fatores da Síndrome Metabólica (SM) ocasiona alterações cardiovasculares, segundo a literatura científica [3, 4, 5, 6, 13, 14]. Entretanto, não foram encontrados artigos científicos que mencionem sistemas de detecção de SM utilizando sinais de Eletrocardiografia (ECG). Logo, a proposta deste trabalho consistiu em suprir essa lacuna científica.

Diante disso, esse trabalho consistiu em implementar diferentes abordagens para detecção de SM utilizando sinais de ECG. Essas abordagens consistem em sistemas classificadores treinados para indicar se um conjunto de características dos sinais de ECG é relacionado a indivíduos saudáveis ou com SM. Para isso, foi desenvolvido um algoritmo para detecção do complexo QRS a fim de realizar a extração de características dos sinais de ECG. Para testar esse algoritmo foi realizada análise do gráfico de Bland e Altman [63]. Essas características foram utilizadas no desenvolvimento de sistemas classificadores, tais como, Máquinas de Vetores Suporte (SVM, do inglês *Support Vector Machine*) e *Robust-Boost*. Além disso, foi utilizada uma técnica para a classificação utilizando aprendizagem profunda, denominada Rede Neural Convolutacional (CNN, do inglês *Convolutional Neural Network*), que realiza a extração de características automaticamente. Ao final, para estimar o desempenho desses sistemas classificadores foi utilizado k-fold, a partir deste foram realizadas medidas estatísticas.

Avaliando os resultados de detecção dos picos das ondas Q, R e S, observou-se que é possível detectar os complexos QRS dos sinais de ECG utilizando o algoritmo implementado. Entretanto, observou-se que na presença de ruídos, gerados por artefatos de movimento, o algoritmo possui seu desempenho prejudicado. Utilizando essa metodologia adotada, os testes de hipóteses indicam que as características extraídas das derivações *DI* e *aVF* dos sinais de ECG, tais como, médias e variâncias das relações e intervalos entre os picos Q, R e S consecutivos e do eixo cardíaco, permitem diferenciar indivíduos do grupo controle com o grupo com SM – $p < 0,01$.

Portanto, foi possível utilizar essas informações extraídas dos sinais de ECG para detecção de SM, a partir do desenvolvimento de classificadores utilizando SVM e *Robust-*

Boost, com sensibilidade de até 96,1%, exatidão de até 97% e especificidade de até 99,9%. Além disso, a CNN para classificação dos sinais de ECG utilizando 12 derivações, permite detectar a SM com resultados estatisticamente significativos, tais como, sensibilidade de até 99,7%, exatidão de até 99,6% e especificidade de até 99,7%.

Comparando esses resultados, observa-se que a CNN obteve resultados estatísticos superiores. Além disso, a extração de características é realizada automaticamente, sem necessidade de utilizar técnicas de processamento de sinais anteriormente. Logo, a CNN é a estratégia mais indicada para ser utilizada como uma ferramenta de detecção de SM por meio de sinais de ECG, em relação às demais abordagens utilizadas neste trabalho.

Em virtude desses aspectos mencionados, em relação aos resultados obtidos ao longo desta dissertação, infere-se que características relacionadas às informações extraídas dos complexos QRS dos sinais de ECG podem diferenciar indivíduos com SM de um grupo sem síndrome. Além disso, é possível detectar a SM a partir dos sinais de Eletrocardiografia com resultados estatisticamente significativos, utilizando diferentes abordagens. Portanto, um eletrocardiograma de 12 derivações pode ser utilizado para avaliar um possível diagnóstico de SM.

Apesar disso, é importante mencionar limitações desse estudo, pois não foram encontradas informações antropométricas de cada participante da pesquisa referentes aos sinais de ECG disponibilizados no banco de dados, tais como, idade, altura, glicose, triglicérido, colesterol, pressão arterial, peso. Portanto, neste estudo, não foi possível analisar a contribuição de cada fator de risco da SM em alterações nos sinais de ECG. Em continuação a esta pesquisa será realizada a aquisição dos sinais de ECG para ter um registro das informações antropométricas de cada indivíduo, e desta forma, incluir, além dos indivíduos saudáveis, mais de um grupo de indivíduos com SM com combinações de diferentes fatores de risco. O intuito é avaliar a contribuição de distintos fatores de risco da SM em alterações provocadas nos sinais de ECG e se isso pode alterar os resultados obtidos pelo sistema classificador.

Referências Bibliográficas

- [1] Sociedade Brasileira de Cardiologia. *Diretriz Brasileira de Diagnóstico e Tratamento da Síndrome Metabólica*, 2005. Arquivos Brasileiros de Cardiologia.
- [2] H. M. Kronenberg, S. Melmed, K. Polonsky, e P. R. Larsen. *Williams: Tratado de Endocrinologia*. Elsevier, 11^a edição, 2010. Tradução de: *Williams textbook of endocrinology*.
- [3] G. Georgiopoulos, C. Tsioufis, D. Tsiachris, K. Dimitriadis, A. Kasiakogias, F. Lagiou, E. Andrikou, I. Ioannidis, E. Hatziagelaki, e D. Tousoulis. Metabolic syndrome, independent of its components, affects adversely cardiovascular morbidity in essential hypertensives. *Atherosclerosis*, 244:66–72, janeiro de 2016.
- [4] A. Kubičková, J. Kozumplík, Z. Nováková, M. Plachý, P. Jurák, e J. Lipoldová. Heart rate variability analysed by Poincaré plot in patients with metabolic syndrome. *Journal of Electrocardiology*, 49(1):23–28, janeiro de 2016.
- [5] S. Jeong, Y. M. Jo, S. O. Shim, Y. J. Choi, e C. H. Youn. A novel model for metabolic syndrome risk quantification based on areal similarity degree. *IEEE Transactions on Biomedical Engineering*, 61(3):665–679, 2014.
- [6] T. W. Elffers, R. de Mutsert, H. J. Lamb, A. C. Maan, P. W. Macfarlane, K. Willems van Dijk, F. R. Rosendaal, J. W. Jukema, e S. Trompet. Association of metabolic syndrome and electrocardiographic markers of subclinical cardiovascular disease. *Diabetology & metabolic syndrome*, 9(1):40, maio de 2017.
- [7] H. K. Pedersen, V. Gudmundsdottir, H. B. Nielsen, T. Hyotylainen, T. Nielsen, B. A. H. Jensen, K. Forslund, F. Hildebrand, E. Prifti, G. Falony, E. Le Chatelier, F. Levenez, J. Doré, I. Mattila, D. R. Plichta, P. Pöhö, L. I. Hellgren, M. Arumugam, S. Sunagawa, S. Vieira-Silva, T. Jørgensen, J. B. Holm, K. Trošt, M. Consortium, K. Kristiansen, S. Brix, J. Raes, J. Wang, T. Hansen, P. Bork, S. Brunak, M. Oresic, S. Dusko Ehrlich, e O. Pedersen. Human gut microbes impact host serum metabolome and insulin sensitivity. *Nature*, 535(7612):376–381, 2016.

- [8] Institute for Health Metrics e Evaluation (IHME). Global burden of disease (GBD), 2018. Em <http://www.healthdata.org>. Último acesso: 31/10/2018.
- [9] W. Lin, H. Zhang, e Y. Zhang. Investigation on cardiovascular risk prediction using physiological parameters. *Computational and mathematical methods in medicine*, 2013:272691, 2013.
- [10] M.S. Thaler. *ECG Essencial: Eletrocardiograma na prática diária*. ARTMED, Porto Alegre, 5ª edição, 2008. Tradução de Jussara N. T. Burnier.
- [11] G. Simonyi. Electrocardiological Features in Obesity : The Benefits of Body Surface Potential Mapping. *Cardiorenal Medicine*, páginas 123–129, 2014.
- [12] E. G. C. de Faria. Sistema de análise de características de sinais de eletrocardiografia para associação com a síndrome metabólica. Dissertação de graduação, em Engenharia Eletrônica na Universidade de Brasília, Campus Gama, DF, 2016.
- [13] E. Severeyn, S. Wong, G. Passariello, J. L. Cevallos, e D. Almeida. Methodology for the study of metabolic syndrome by heart rate variability and insulin sensitivity. *Revista Brasileira de Engenharia Biomédica*, 28:272 – 277, setembro de 2012.
- [14] G. Perpi, Student Member, E. Severeyn, M. Altuve, e S. Wong. Classification of Metabolic Syndrome Subjects and Marathon Runners with the k -means Algorithm using Heart Rate Variability Features. *XXI Symposium on Signal Processing, Images and Artificial Vision (STSIVA)*, páginas 1–6, 2016.
- [15] Scott M. Grundy. Pre-diabetes, metabolic syndrome, and cardiovascular risk. *Journal of the American College of Cardiology*, 59(7):635–643, 2012.
- [16] Tamara B. Horwich e Gregg C. Fonarow. Glucose, Obesity, Metabolic Syndrome, and Diabetes. Relevance to Incidence of Heart Failure. *Journal of the American College of Cardiology*, 55(4):283–293, 2010.
- [17] Paul L Hess, Hussein R Al-khalidi, Daniel J Friedman, Hillary Mulder, Anna Kucharska-newton, Wayne R Rosamond, Renato D Lopes, Bernard J Gersh, Daniel B Mark, Lesley H Curtis, Wendy S Post, Ronald J Prineas, Nona Sotoodehnia, e Sana M Al-khatib. The Metabolic Syndrome and Risk of Sudden Cardiac Death: The Atherosclerosis Risk in Communities Study. *Journal of the American Heart Association*, 6(8), agosto de 2017.
- [18] Pranav Rajpurkar, Codie Bourn, Andrew Y Ng, Pranavsr Cs, Stanford Edu, A N G Cs, e Stanford Edu. Cardiologist-Level Arrhythmia Detection with Convolutional Neural Networks. *CoRR*, julho de 2017.

- [19] Zhaohan Xiong, Martin Stiles, e Jichao Zhao. Robust ECG Signal Classification for the Detection of Atrial Fibrillation Using Novel Neural Networks. *2017 Computing in Cardiology (CinC)*, 44:1–4, setembro de 2017.
- [20] Zhongdi Liu, XiangAo Meng, Jiajia Cui, Zhipei Huang, e Jiankang Wu. Automatic Identification of Abnormalities in 12-Lead ECGs Using Expert Features and Convolutional Neural Networks. *2018 International Conference on Sensor Networks and Signal Processing (SNSP)*, 877:163–167, 2018.
- [21] Yufa Xia e Yaoqin Xie. A novel wearable electrocardiogram classification system using convolutional neural networks and active learning. *IEEE Access*, julho de:1–13, 2019.
- [22] Zhaohan Xiong, Martyn P. Nash, Elizabeth Cheng, Vadim V. Fedorov, Martin K. Stiles, e Jichao Zhao. ECG signal classification for the detection of cardiac arrhythmias using a convolutional recurrent neural network. *Physiological Measurement*, 39(9), 2018.
- [23] R. Merletti e P.A. Parker. *Electromyography: Physiology, Engineering, and Non-Invasive Applications*. IEEE Press Series on Biomedical Engineering. Wiley, 2004.
- [24] E. V. Garcia. *Processamento de Sinais usando Wavelets para Caracterização da Repolarização Ventricular durante Hipoglicemia*. Tese de Doutorado, Universidade Federal de Santa Catarina, Florianópolis, 2005.
- [25] B. Koeppen. *Berne e Levy Fisiologia*. Elsevier Brasil, 2011.
- [26] G.P. Goldwasser. *Eletrocardiograma Orientado Para O Clínico*. RUBIO, 3ª edição, 2009.
- [27] J. Hallake. *Eletrocardiografia*. Medsi, Rio de Janeiro, 3ª edição, 2004.
- [28] Ashly Elizabeth Vincent. A Survey on Approaches for ECG Signal Analysis with Focus to Feature Extraction and Classification. *International Conference on Inventive Communication and Computational Technologies (ICICCT)*, páginas 140–144, 2017.
- [29] C.F. Camm e A.J. Camm. *Clinical Guide to Cardiology*. Clinical Guides. Wiley, 2016.
- [30] Willis J. Tompkins. *Biomedical Digital Signal Processing: C-language Examples and Laboratory Experiments for the IBM PC*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1993.

- [31] Ari Timerman, M. C. Bertolami, e J. F. M. Ferreira. *Manual de Cardiologia*. Atheneu, 2012.
- [32] Jiapu Pan e Willis J. Tompkins. A Real-Time QRS Detection Algorithm. *IEEE transactions on Biomedical Engineering*, BME-32(3):230–236, 1985.
- [33] Zhi-Hua Zhou. *Ensemble Methods: Foundations and Algorithms*. Chapman & Hall/CRC, 1ª edição, 2012.
- [34] Chih-wei Hsu, Chih-chung Chang, e Chih-jen Lin. A Practical Guide to Support Vector Classification. Technical Report 1, Department of Computer Science, National Taiwan University, 2010.
- [35] Yann LeCun, Y Bengio, e Geoffrey Hinton. Deep Learning. *Nature*, 521:436–444, maio de 2015.
- [36] Runze Yu, Yingguo Gao, Xiaohui Duan, Tiangang Zhu, Zhilong Wang, e Bingli Jiao. QRS Detection and Measurement Method of ECG Paper Based on Convolutional Neural Networks. *Conference proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, páginas 4636–4639, julho de 2018.
- [37] Vincentius Timothy, Ary Setijadi Prihatmanto, e Kyung-hyune Rhee. Data Preparation Step for Automated Diagnosis based on HRV Analysis and Machine Learning. *2016 6th International Conference on System Engineering and Technology (ICSET)*, páginas 142–148, 2016.
- [38] W. Li e J. Li. Local deep field for electrocardiogram beat classification. *IEEE Sensors Journal*, 18(4):1656–1664, fevereiro de 2018.
- [39] Jinlong Ji, Student Member, Xuhui Chen, e Student Member. A deep multi-task learning approach for ECG data analysis. In *2018 IEEE EMBS International Conference on Biomedical Health Informatics (BHI)*, páginas 124–127, março de 2018.
- [40] Serkan Kiranyaz, Turker Ince, Ridha Hamila, e Moncef Gabbouj. Convolutional Neural Networks for patient-specific ECG classification. In *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, páginas 2608–2611, agosto de 2015.
- [41] T. Yang, L. Yu, Q. Jin, L. Wu, e B. He. Localization of Origins of Premature Ventricular Contraction by Means of Convolutional Neural Network From 12-Lead ECG. *IEEE Transactions on Biomedical Engineering*, 65(7):1662–1671, julho de 2018.

- [42] Vykintas Maknickas, Algirdas Maknickas, e Tesonet Llc. Atrial Fibrillation Classification Using QRS Complex Features and LSTM. *Computing in Cardiology*, 44:1–4, 2017.
- [43] Rishikesan Kamaleswaran, Ruhi Mahajan, e Oguz Akbilgic. A robust deep convolutional neural network for the classification of abnormal cardiac rhythm using varying length single lead electrocardiogram. In *Computing in Cardiology 2017*, volume 39, janeiro 2018.
- [44] G. Clifford, C. Liu, B. Moody, I. Silva, Q. Li, A. Johnson, e R. Mark. AF Classification from a Short Single Lead ECG Recording: the PhysioNet Computing in Cardiology Challenge. *Computing in Cardiology*, 44(In Press):1–4, 2017.
- [45] Yong Xia, Naren Wulan, Kuanquan Wang, e Henggui Zhang. Atrial Fibrillation Detection using Stationary Wavelet Transform and Deep Learning. *Computing in Cardiology 2017*, 44:2–5, setembro de 2017.
- [46] Joanna Rymko, Mateusz Solinski, Anna Perka, Jacek Rosinski, e Michal Lepek. Classification of Atrial Fibrillation in Short-term ECG Recordings Using a Machine Learning Approach and Hybrid QRS Detection. *Computing in Cardiology 2017*, 44:1–4, 2017.
- [47] Shweta H. Jambukia, Vipul K. Dabhi, e Harshadkumar B. Prajapati. Classification of ECG signals using machine learning techniques: A survey. In *2015 International Conference on Advances in Computer Engineering and Applications*, volume 4, páginas 714–721. IEEE, março de 2015.
- [48] Yan Ma, Ping Huei Tseng, Andrew Ahn, Ming Shiang Wu, Yi Lwun Ho, Ming Fong Chen, e Chung Kang Peng. Cardiac Autonomic Alteration and Metabolic Syndrome: An Ambulatory ECG-based Study in A General Population. *Scientific Reports*, 7:1–9, março de 2017.
- [49] S. Tani, H. Narazaki, M. Iwata, R. Haraguchi, S. Kuwata, Hi. Inada, e K. Nakazawa. Feasibility for the enhancement of an online support system for persons with metabolic syndrome, aimed at applications for ischemic heart disease and heart failure. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, (Table I):1190–1193, 2013.
- [50] Jared W Magnani, Elaine M. Hylek, e Caroline M. Apovian. Obesity begets Atrial Fibrillation: A Contemporary Summary Jared. *Circulation*, julho de 2013.
- [51] Stuart Frank, Jerry A. Colliver, e Arthur Frank. The electrocardiogram in obesity: Statistical analysis of 1,029 patients. *Journal of the American College of Cardiology*, 7(2):295–299, 1986.

- [52] Guo Zhe Sun, Yang Li, Xing Hu Zhou, Xiao Fan Guo, Xin Gang Zhang, Li Qiang Zheng, Yuan Li, Yun Di Jiao, e Ying Xian Sun. Association between obesity and ECG variables in children and adolescents: A cross-sectional study. *Experimental and Therapeutic Medicine*, 6(6):1455–1462, 2013.
- [53] M. Altuve, E. Severeyn, e S. Wong. Unsupervised subjects classification using insulin and glucose data for insulin resistance assessment. In *2015 20th Symposium on Signal Processing, Images and Computer Vision (STSIVA)*, páginas 1–7. IEEE, setembro de 2015.
- [54] B. Bagheri. Uric Acid and Coronary Artery Disease, Two Sides of a Single Coin: A Determinant of Antioxidant System or a Factor in Metabolic Syndrome. *Journal of Clinical and Diagnostic Research*, 10(2):27–31, 2016.
- [55] Georgeta D Vaidean, Marta Manczuk, Jared W Magnani, Florham Park, The Cancer Epidemiology, e Prevention Division. Atrial Electrocardiography in Obesity and Hypertension: Clinical Insights from the Polish-Norwegian Study (PONS). *Obesity (Silver Spring)*, 24(12):2608–2614, 2017.
- [56] Sinan Akinci e C Davran. Which is responsible for cardiac autonomic dysfunction in non-diabetic patients with metabolic syndrome : Prediabetes or the syndrome itself? *Diabetes & Metabolic Syndrome : Clinical Research & Reviews*, 2016.
- [57] R. Musale e A. N. Paithane. Design and develop an algorithm for a diabetic detection using ECG signal. In *2017 International Conference on Computing Methodologies and Communication (ICCMC)*, páginas 961–966, julho de 2017.
- [58] Tarak Das, Arijit Ghosh, Sayanti Guha, e Piyali Basak. Early detection of diabetes based on skin impedance spectrogram and heart rate variability noninvasively. In *2017 1st International Conference on Electronics, Materials Engineering and Nano-Technology (IEMENTech)*, páginas 1–5. IEEE, abril de 2017.
- [59] I. Cavallari, C. P Cannon, E. Braunwald, E. L Goodrich, K. Im, M. A. Lukas, e M. L O’Donoghue. Metabolic syndrome and the risk of adverse cardiovascular events after an acute coronary syndrome. *European Journal of Preventive Cardiology*, 67(13):204748731876389, março de 2018.
- [60] Phyto Phyto San, Sai Ho Ling, e Hung T. Nguyen. Deep learning framework for detection of hypoglycemic episodes in children with type 1 diabetes. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, 2016(Cd):3503–3506, agosto de 2016.

- [61] Hooman Sedghamiz. Matlab implementation of Pan Tompkins ECG QRS detector, março de 2014. Em https://www.researchgate.net/publication/313673153_Matlab_Implementation_of_Pan_Tompkins_ECG_QRS_detector. Último acesso: 22/07/2018.
- [62] Vika Octaviani, Arief Kurniawan, Yoyon Kusnendar Suprpto, e Ahmad Zaini. Alerting System for Sport Activity Based on ECG Signals using Proportional Integral Derivative. *Proc. EECSI 2017*, páginas 19–21, setembro de 2017.
- [63] J. Martin Bland e Douglas G Altman. Statistical Methods for Assessing Agreement Between Two Methods of Clinical Measurement. *The Lancet*, 327(8476):307–310, 1986.
- [64] David Spodick, Mary Frisella, e Sirin Apiyassawat. QRS Axis Validation in Clinical Electrocardiography. *The American journal of cardiology*, 101:268–9, fevereiro de 2008.
- [65] Mats Halldin, Kerstin Brismar, Per Fahlstadius, Max Vikström, Ulf De Faire, e Mai Lis Hellénus. The metabolic syndrome and ECG detected left ventricular hypertrophy - Influences from IGF-1 and IGF-binding protein-1. *PLoS ONE*, 9(12):1–13, 2014.
- [66] Marijana Tadic, Branislava Ivanovic, Nada Kostic, Dragan Simic, Danica Matic, e Vera Celic. Metabolic syndrome and left ventricular function: is the number of criteria actually important? *Medical science monitor : international medical journal of experimental and clinical research*, 18(5), maio de 2012.
- [67] Göksu Bozdereli Berikol, Oktay Yildiz, e Ä°. Türkay Özcan. Diagnosis of Acute Coronary Syndrome with a Support Vector Machine. *Journal of Medical Systems*, 40(4):84, abril de 2016.
- [68] Angshul Majumdar e Rabab Ward. Robust greedy deep dictionary learning for ECG arrhythmia classification. In *2017 International Joint Conference on Neural Networks (IJCNN)*, páginas 4400–4407. IEEE, maio de 2017.
- [69] B Pourbabae, M J Roshtkhari, e K Khorasani. Deep Convolutional Neural Networks and Learning ECG Features for Screening Paroxysmal Atrial Fibrillation Patients. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, páginas 1–10, 2017.
- [70] Serkan Kiranyaz, Turker Ince, e Moncef Gabbouj. Real-Time Patient-Specific ECG Classification by 1-D Convolutional Neural Networks. *IEEE Transactions on Biomedical Engineering*, 63(3):664–675, março de 2016.

- [71] GB Moody, WE Muldrow, e RG Mark. A noise stress test for arrhythmia detectors. *Computers in Cardiology*, 11:381–384, 1984.
- [72] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. Ch. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, e H. E. Stanley. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation*, 101(23):e215–e220, junho de 2000.
- [73] C. A. Ledezma, E. Severeyn, G. Perpiñán, M. Altuve, e S. Wong. A new on-line electrocardiographic records database and computer routines for data analysis. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, páginas 2738–2741, agosto de 2014.
- [74] Cardiosoft. Gravador digital de ECG de 12 derivações e software para testes de repouso e de exercícios. Pesquisa avançada em cardiologia e desenvolvimento de aplicações inovadoras - do inglês, *Enhanced Cardiology Research and Development of Innovative Applications (Encardia)*. Em <http://www.cardiosoft.com>. Último acesso: 20/06/2018.
- [75] Ron Kohavi. A study of cross-validation and bootstrap for accuracy estimation and model selection. *Proceedings of the 14th international joint conference on Artificial intelligence - Volume 2*, 2:1137–1143, 1995.
- [76] Payam Refaeilzadeh, Lei Tang, e Huan Liu. Cross validation. *Encyclopedia of Database Systems*, páginas 532–538, janeiro de 2009.
- [77] Klein Ran. Bland-Altman and Correlation Plot, Março de 2019. Em <https://www.mathworks.com/matlabcentral/fileexchange/45049-bland-altman-and-correlation-plot>. Último acesso: 17/03/2019.
- [78] Steven R. Steinhubl e Eric J. Topol. Moving from Digitalization to Digitization in Cardiovascular Care Why Is it Important, and What Could it Mean for Patients and Providers? *Journal of the American College of Cardiology*, 66(13):1489–1496, 2015.

APÊNDICE A

PRINCIPAIS INFORMAÇÕES EXTRAÍDAS DOS SINAIS DE ELETROCARDIOGRAFIA

Um sinal de Eletrocardiografia (ECG), que está mostrado na Figura A.1, é constituído principalmente pela onda P, pelo complexo QRS e pela onda T. A onda P representa o período de despolarização atrial, o complexo QRS representa a despolarização ventricular, e a onda T representa a repolarização ventricular.

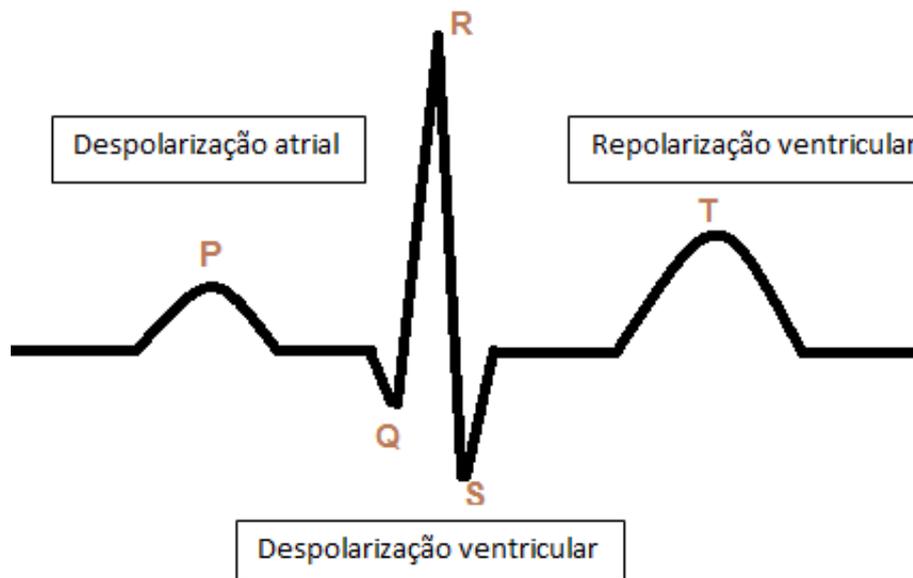


Figura A.1. Representação gráfica de um sinal de eletrocardiografia indicando as fases de despolarização e repolarização. Fonte própria.

As diferenças entre regiões avaliadas no domínio do tempo (eixo horizontal) são representadas por linhas que caracterizam um segmento ou um intervalo. Um segmento é uma linha reta que conecta duas ondas, já o intervalo é uma linha reta que contém pelo menos uma onda mais a linha reta conectada [10]. A Figura A.2 mostra exemplos de segmentos e intervalos obtidos a partir do sinal de ECG.

Esses segmentos e intervalos medem a duração dos seguintes eventos:

- Intervalo PR: início da despolarização atrial até o início da despolarização ventricular;
- Segmento PR: final da despolarização atrial até o início da despolarização ventricular;

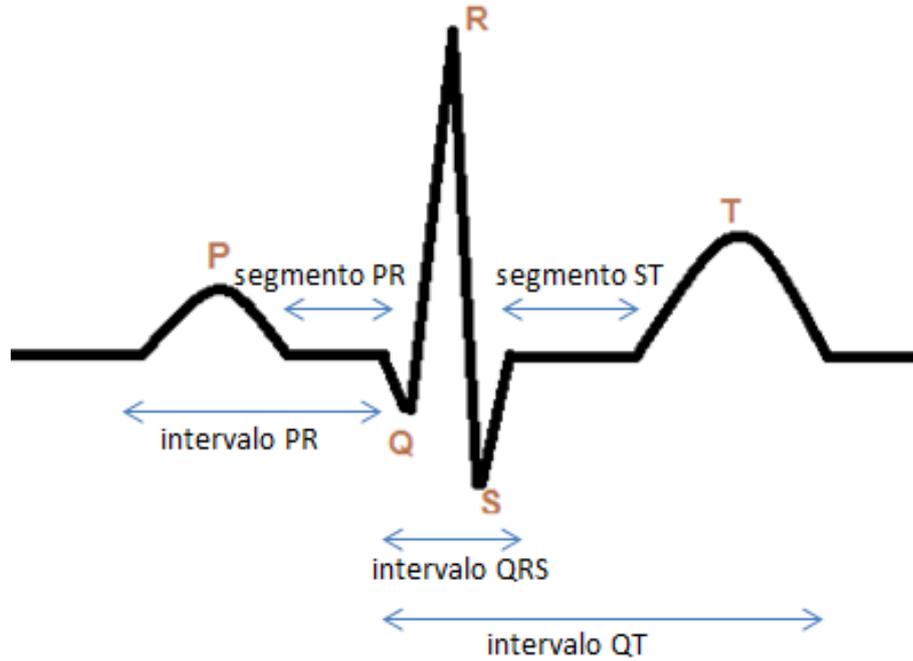


Figura A.2. Representação gráfica de um sinal de eletrocardiografia indicando seus principais segmentos e intervalos. Adaptado de [10].

- c. Intervalo QRS: despolarização ventricular;
- d. Intervalo QT: início da despolarização ventricular até o final da repolarização ventricular; e
- e. Segmento ST: final da despolarização ventricular até o início da repolarização ventricular [10].

Além desses intervalos e segmentos, o intervalo RR consiste na distância entre dois picos R e o intervalo QTc consiste no intervalo QT corrigido para frequência, que utilizando a fórmula de Bazett é calculado conforme a Equação 5.1.

$$QTc = \frac{QT}{\sqrt{RR(s)}}, \quad (5.1)$$

sendo QT o intervalo QT medido e RR o intervalo entre os picos R dos sinais de ECG [27].

EIXO CARDÍACO

O termo eixo cardíaco possui como definição a direção do vetor elétrico médio que registra a atividade elétrica do coração. O eixo cardíaco pode estar relacionado a cada um dos fenômenos elétricos – onda P, complexo QRS ou onda T; entretanto o principal eixo analisado é o do QRS que registra o fenômeno de despolarização, que é o momento de registro no ECG da maior parte da força elétrica, pois o ventrículo esquerdo consiste na maior parte do músculo cardíaco [10].

A análise do eixo cardíaco ocorre avaliando-se o ciclo trigonométrico da Figura B.1, onde, no eixo frontal, cada uma das derivações possuem uma orientação diferente [10]. A derivação *DI* é a referência (eixo de grau 0) do eixo cardíaco, as demais derivações são deslocadas a cada 30 graus conforme a Figura B.1.

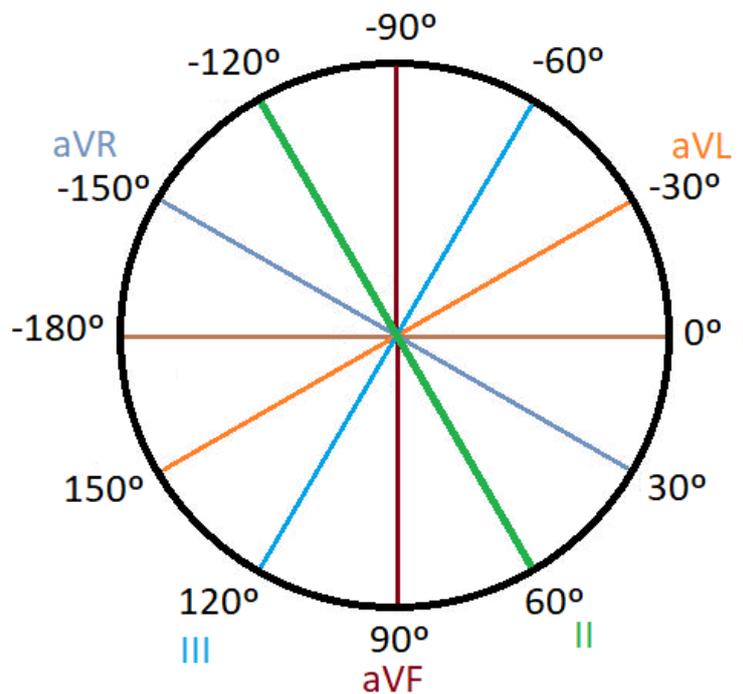


Figura B.1. Representação gráfica para análise do eixo cardíaco com respeito as derivações *DI*, *DII*, *DIII*, *aVR*, *aVL* e *aVF* do plano frontal. Adaptado de [26, 27].

Para calcular o eixo cardíaco é analisada a predominância do complexo QRS em cada derivação. Caso o tamanho vertical da onda R seja predominante, o vetor possui sentido positivo nesta derivação, se o S for predominante, o vetor possui sentido negativo.

Um dos métodos de análise do eixo cardíaco utilizado por cardiologistas e clínicos é o método dos quadrantes que consiste em avaliar a região do vetor conforme apresentado na Figura B.2. Quando o vetor analisado estiver na região entre -30° e $+90^\circ$ o ECG é considerado normal; entre -90° até -30° é considerado desvio para a esquerda; entre -180° e $+90^\circ$ é considerado um desvio para direita. Na região de 180° a -90° é considerado um desvio extremo do eixo para a direita [10, 27].

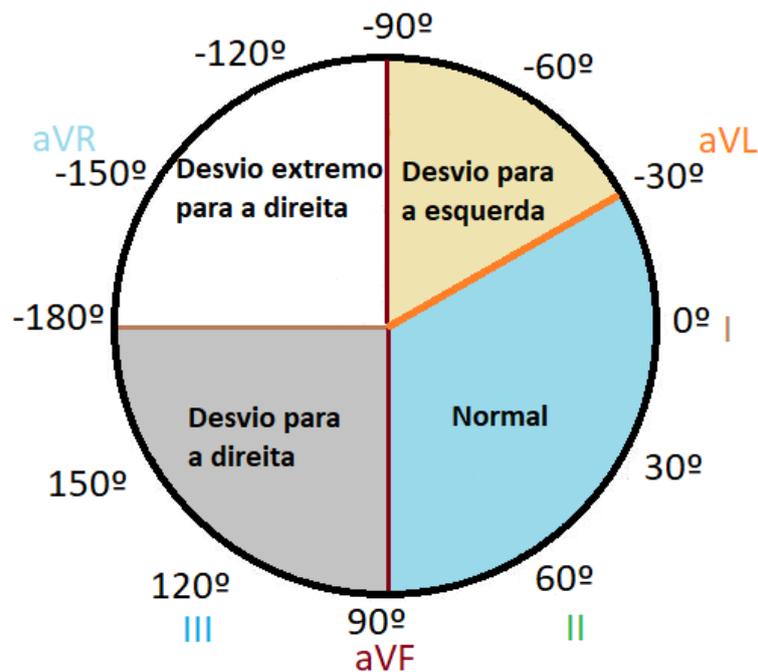


Figura B.2. Representação das regiões consideradas de desvio do eixo e a região considerada como normal depois da identificação do eixo cardíaco. Adaptado de [10, 29].

Analisando a Figura B.2 é possível notar que as derivações *DI*, *aVF* e *DII* são suficientes para determinar em qual quadrante está localizado o eixo cardíaco. Sabendo-se que o eixo cardíaco não se encontra na região de -30° a 0° , apenas as derivações *DI* e *aVF* são suficientes para determinar o quadrante do eixo cardíaco [29].

CARACTERÍSTICAS EXTRAÍDAS DOS SINAIS DE ELETROCARDIOGRAFIA

Tabela C.1. Legenda das características extraídas dos sinais de ECG apresentadas na Tabela 4.6.

Característica extraída do sinal de ECG	
1	média dos intervalos RR da derivação <i>DI</i>
2	média dos intervalos RS da derivação <i>DI</i>
3	média dos intervalos QS da derivação <i>DI</i>
4	média dos intervalos QR da derivação <i>DI</i>
5	média do eixo cardíaco, calculado com base na derivação <i>DI</i> e <i>aVF</i>
6	média da relação entre R e Q da derivação <i>DI</i>
7	média da relação entre R e S da derivação <i>DI</i>
8	média da amplitude de R da derivação <i>DI</i>
9	variância dos intervalos RR da derivação <i>DI</i>
10	variância dos intervalos RS da derivação <i>DI</i>
11	variância dos intervalos QS da derivação <i>DI</i>
12	variância dos intervalos QR da derivação <i>DI</i>
13	variância do eixo cardíaco, calculado com base na derivação <i>DI</i> e <i>aVF</i>
14	variância da relação entre R e Q da derivação <i>DI</i>
15	variância da relação entre R e S da derivação <i>DI</i>
16	variância da amplitude de R da derivação <i>DI</i>
17	média dos intervalos RR da derivação <i>aVF</i>
18	média dos intervalos RS da derivação <i>aVF</i>
19	média dos intervalos QS da derivação <i>aVF</i>
20	média dos intervalos QR da derivação <i>aVF</i>
21	média da relação entre R e Q da derivação <i>aVF</i>
22	média da relação entre R e S da derivação <i>aVF</i>
23	média da amplitude de R da derivação <i>aVF</i>
24	variância dos intervalos RR da derivação <i>aVF</i>
25	variância dos intervalos RS da derivação <i>aVF</i>
26	variância dos intervalos QS da derivação <i>aVF</i>
27	variância dos intervalos QR da derivação <i>aVF</i>
28	variância da relação entre R e Q da derivação <i>aVF</i>
29	variância da relação entre R e S da derivação <i>aVF</i>
30	variância da amplitude de R da derivação <i>aVF</i>