

**QUENAZ BEZERRA SOARES**

**Modelos de aprendizado profundo de baixo custo  
computacional para classificação de ritmo cardíaco**

São Paulo  
2024

**QUENAZ BEZERRA SOARES**

**Modelos de aprendizado profundo de baixo custo computacional para classificação de ritmo cardíaco**

Versão Corrigida

Dissertação apresentada à Escola Politécnica da Universidade de São Paulo para obtenção do Título de Mestre em Ciências.

Área de Concentração:  
Engenharia Biomédica

Orientador:  
Prof. Dr. Marco Antonio Gutierrez

São Paulo  
2024

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Este exemplar foi revisado e corrigido em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

São Paulo, 12 de Junho de 2024

Assinatura do autor: Quenaz Bezerra Soares

Assinatura do orientador: [Assinatura]

#### Catálogo-na-publicação

Soares, Quenaz Bezerra

Modelos de aprendizado profundo de baixo custo computacional para classificação de ritmo cardíaco / Q. B. Soares -- versão corr. -- São Paulo, 2024. 115 p.

Dissertação (Mestrado) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Telecomunicações e Controle.

1.Fibrilação Atrial 2.Arritmia 3.Aprendizagem Profunda I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Telecomunicações e Controle II.t.

À minha família e amigos.

# AGRADECIMENTOS

Quero expressar meus sinceros agradecimentos a várias pessoas e instituições que contribuíram para a conclusão deste trabalho.

Primeiramente, ao meu orientador, Dr. Marco, pela orientação valiosa, apoio constante e paciência ao longo deste processo. Suas orientações e insights foram fundamentais para o desenvolvimento desta dissertação.

À minha família, por seu amor incondicional, apoio emocional e encorajamento ao longo de toda a minha jornada acadêmica. Sem vocês, nada disso seria possível.

À minha querida namorada, Letícia, pela compreensão, apoio e incentivo constantes. Sua presença tornou os desafios mais leves e as conquistas mais significativas.

Agradeço também aos meus colegas de trabalho, que compartilharam ideias, experiências e desafios acadêmicos, contribuindo para o meu crescimento e aprendizado.

À equipe do InCor (Instituto do Coração), por disponibilizar recursos e ambiente propício à pesquisa, bem como a todos os profissionais que colaboraram de alguma forma com este trabalho.

Agradeço ao CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) e à USP (Universidade de São Paulo) pelo suporte financeiro e acadêmico fornecido durante este estudo.

Por fim, agradeço a todos os amigos, familiares e mentores que, de alguma forma, contribuíram para esta jornada. Seu apoio foi fundamental para meu crescimento pessoal e acadêmico.

Este trabalho não teria sido possível sem o suporte e encorajamento de todos vocês. Muito obrigado.

# RESUMO

A Fibrilação Atrial (FA) é a arritmia cardíaca mais comum no mundo, com prevalência superior a 5% na população acima de 70 anos. Possíveis consequências da FA incluem eventos tromboembólicos e insuficiência cardíaca. O monitoramento contínuo de pacientes de risco, como os pós-operatórios de cirurgia cardíaca, é uma importante ferramenta para detectar essa arritmia, que pode se apresentar de forma intermitente. Tal monitoramento pode ser realizado com eletrocardiógrafos vestíveis, associados a detectores automáticos de eventos de FA. Esse monitoramento é limitado pela complexidade dos detectores utilizados, sendo que os mais robustos são baseados em aprendizado profundo e geralmente requerem alto poder computacional. Este trabalho visa o desenvolvimento de arquiteturas de aprendizado profundo, com requisitos computacionais compatíveis com plataformas móveis, ou mesmo microcontroladores, mantendo desempenho de classificação comparável a modelos tradicionais mais complexos. Com base nas arquiteturas de redes neurais convolucionais VGGNet e ResNet, foi proposto o uso de convoluções separáveis, menos filtros, *poolings* globais e menos unidades nas camadas densas, dando origem às arquiteturas LiteVGG e LiteResNet. Uma arquitetura ainda mais leve também foi proposta, usando convoluções em profundidade (*depthwise*) e menos camadas, chamada DWNet. As arquiteturas originais VGGNet, ResNet e um conjunto de arquiteturas encontradas na literatura foram utilizadas para comparação com as arquiteturas propostas. A base de dados publicada para o *PhysioNet Challenge 2021* foi utilizada para esta comparação, reestruturando os diagnósticos para classificação binária (FA e não FA), multiclasse (FA, ritmo sinusal e outros) e multilabel (FA, ritmo sinusal normal, taquicardia sinusal e bradicardia sinusal). A comparação das métricas de classificação foi realizada por meio de validação cruzada Monte Carlo, e uma comparação de custo computacional foi realizada em termos de número de operações de ponto flutuante (FLOP), número de parâmetros, tempo de inferência e memória utilizada. Os modelos propostos apresentaram redução no número de FLOPs e parâmetros em cerca de 2 a 4 ordens de grandeza quando comparados aos modelos utilizados como base, além de uma redução de 1 a 2 ordens de grandeza no tempo de inferência. Para a arquitetura VGGNet, essa redução na complexidade resultou em um modelo LiteVGG-11 mais leve e superior em termos de métricas que seu equivalente VGG-11. Este foi o melhor modelo encontrado para reduzir falsos positivos em todas as estratégias de classificação de FA testadas. A arquitetura LiteResNet apresentou métricas de classificação, por vezes, sem diferença estatística em relação ao ResNet. O modelo DWNet possui a menor complexidade computacional, mas com a desvantagem de pior acurácia. A estratégia adotada para reduzir o custo computacional resultou em modelos mais leves e com desempenhos de classificação semelhantes ou melhores que os modelos tradicionais mais robustos. Com esses modelos, seria possível implementar sistemas de monitoramento de ritmo cardíaco que poderiam ser embarcados em dispositivos móveis ou até mesmo em microcontroladores.

**Palavras-Chave** – Fibrilação Atrial, Arritmia, Aprendizagem Profunda.

# ABSTRACT

Atrial Fibrillation (AF) is the most common cardiac arrhythmia in the world, with a prevalence greater than 5% in the population over 70 years of age. Possible consequences of AF include thromboembolic events and heart failure. Continuous monitoring of high-risk patients, such as those after cardiac surgery, is an important tool to detect this arrhythmia, which may present intermittently. Such monitoring can be performed using wearable electrocardiographs, associated with automatic AF event detectors. This monitoring is limited by the complexity of the detectors used, the most robust ones being based on deep learning and generally requiring high computational power. This work aims at the development of deep learning architectures, with computational requirements compatible with mobile platforms, or even microcontrollers, while maintaining classification performance comparable to more complex traditional models. Based on the VGGNet and ResNet convolutional neural network architectures, it was proposed to use separable convolutions, fewer filters, global poolings and fewer units in the dense layers, giving rise to LiteVGG and LiteResNet architectures. An even lighter architecture was also proposed, using depthwise convolutions and fewer layers, called DWNNet. The original architectures VGGNet, ResNet and a set of architectures found in the literature were used for comparison with the proposed architectures. The database published for the PhysioNet Challenge 2021 was used for this comparison, restratifying the diagnoses for binary classification (AF and non-AF), multiclass (AF, sinus rhythm and others) and multilabel (AF, normal sinus rhythm, sinus tachycardia and sinus bradycardia). Comparison of classification metrics was performed using Monte Carlo cross-validation, and a computational cost comparison was performed in terms of number of floating point operations, number of parameters, inference time and memory used. The proposed models showed a reduction in the number of FLOPs and parameters by about 2 to 4 orders of magnitude when compared to the models used as baseline, in addition to a reduction of 1 to 2 orders of magnitude in the inference time. For the VGGNet architecture, this reduction in complexity resulted in a LiteVGG-11 model that is lighter and better in terms of metrics than its VGG-11 counterpart. This was the best model found to reduce false positives in all AF classification schemes tested. LiteResNet architecture presented classification metrics with no statistical difference in relation to ResNet for some comparisons. The DWNNet model has the lowest computational complexity, but with the disadvantage of worse accuracy. The strategy adopted to reduce the computational cost resulted in lighter models with similar or better classification performances than the more robust traditional models. With these models, it would be possible to implement heart rate monitoring systems that could be embedded in mobile devices or even microcontrollers.

**Keywords** – Atrial Fibrillation, Arrhythmia, Deep Learning.

# LISTA DE FIGURAS

1	Percentual de mortes causadas por Fibrilação Atrial ao longo dos anos. Área sombreada indica intervalo de confiança. . . . .	17
2	Prevalência de Fibrilação Atrial em 2019, por grupos de idade. . . . .	18
3	Procedimento de aquisição de ECG de repouso. . . . .	23
4	Formato de onda de um ECG típico. . . . .	23
5	Exemplo de um ECG com ritmo normal. . . . .	24
6	Diagrama do coração humano. . . . .	25
7	Fases do potencial de ação de células cardíacas. . . . .	27
8	Sistema de condução de potenciais de ação do coração. . . . .	29
9	Instantes aproximados do início da despolarização dos cardiomiócitos em diferentes partes do coração (em segundos após a despolarização inicial do nó sinoatrial (SA)). . . . .	30
10	Representação esquemática dos potenciais de ação em diferentes regiões do coração e suas correlações com o ECG. . . . .	30
11	Eventos do ciclo cardíaco, mostrando pressão no ventrículo esquerdo, pressão no átrio esquerdo, pressão na aorta, volume do ventrículo esquerdo, ECG e fonocardiograma. . . . .	31
12	Derivação II do ECG de um paciente com ritmo RSN. . . . .	32
13	Derivação II do ECG de um paciente com ritmo TS. . . . .	32
14	Derivação II do ECG de um paciente com ritmo BS. . . . .	33
15	Vias de propagação dos impulsos (A) na fibrilação atrial e (B) no flutter atrial. . . . .	34
16	Derivação II do ECG de um paciente com ritmo FA. . . . .	34
17	Derivação II do ECG de um paciente com ritmo FLU. . . . .	35
18	Componentes do sinal de ECG mensurado. . . . .	37



19	Circuito típico de um eletrocardiograma de uma derivação. . . . .	37
20	Posicionamento dos eletrodos na configuração de ECG de 12 derivações. . .	38
21	Triângulo de Einthoven obtido com os eletrodos dos membros. . . . .	39
22	Diagrama de Venn associando inteligência artificial, aprendizado de máquina e <i>deep learning</i> . . . . .	42
23	Exemplo de árvore de decisão. . . . .	43
24	Exemplos de funções de ativação. . . . .	45
25	Diagrama de uma rede neural artificial do tipo perceptron multicamada. .	45
26	Visualização da operação de convolução 1D sem padding. . . . .	47
27	Visualização da operação de convolução 1D para múltiplos filtros e entrada com múltiplos canais. . . . .	48
28	Exemplo de operações de AVGPool e MaxPool, considerando <i>stride</i> e janela de tamanho 2. . . . .	49
29	Exemplo de operações de <i>Global Average Pooling</i> (GAP) e <i>Global Max Pooling</i> (GMP). . . . .	50
30	Esquemático de alto nível da metodologia do trabalho. . . . .	52
31	Distribuição das idades dos pacientes por gênero e ritmo. Linhas pontilhadas delimitam o intervalo interquartil de cada curva de distribuição. . . . .	55
32	Diagrama de blocos do pré-processamento adotado. . . . .	58
33	(A) Sinal original de Eletrocardiograma (ECG) e sinais pré-processados: (B) ECG limpo, (C) sinal suavizado e (D) posição dos complexos QRS. . .	59
34	Bloco residual genérico. . . . .	60
35	Comparação entre a arquitetura dos modelos VGG-11 e LiteVGG-11. . . .	64
36	Arquitetura do modelo DWNet. . . . .	67
37	Matriz de confusão. . . . .	69
38	Curva ROC. . . . .	70
39	Relação de custo computacional <i>versus</i> taxa de falsos negativos na classificação binária. . . . .	81

40	Impacto do pré-processamento nas métricas da classificação binária. Barras representam valor médio e traços indicam o intervalo de confiança a 95%. . . . .	87
41	Acurácia da classificação binária de acordo com a Relação Sinal-Ruído (SNR) simulada. Região sombreada indica o intervalo de confiança a 95%. . . . .	88
42	Diagrama das arquiteturas ResNet-18 e ResNet-34. . . . .	102
43	Diagrama das arquiteturas LiteResNet-18 e LiteResNet-34. . . . .	103

# LISTA DE TABELAS

1	Percentual de mortes relacionadas a doenças cardiovasculares. . . . .	16
2	Distribuição dos exames por classes de ritmos de interesse. . . . .	54
3	Configurações convencionais de camadas convolucionais em cada bloco da arquitetura VGGNet. Camadas convolucionais estão nomeadas como $convx-y$ , sendo $x$ o tamanho do kernel e $y$ o número de filtros da camada. . . . .	63
4	Comparação entre o número de operações de ponto flutuante (FLOP) e parâmetros dos modelos VGG-11 e LiteVGG-11, bem como das adaptações intermediárias. Linhas indicadas por uma seta para baixo ( $\downarrow$ ) indicam a redução percentual entre o modelo anterior e o posterior. . . . .	65
5	Custo computacional de todos os modelos testados ( <i>baselines</i> e propostos). Asterisco (*) indica modelos avaliados com frequência de amostragem de 200 Hz. . . . .	77
6	Requisitos computacionais mínimos para execução dos modelos em microcontrolador ARM. * indica modelo com frequência de amostragem de 200 Hz. . . . .	79
7	Métricas obtidas para o experimento robusto de classificação binária. . . . .	80
8	Métricas obtidas para a classe <i>Fibrilação Atrial (FA)</i> no experimento robusto de classificação multiclasse. . . . .	82
9	Métricas obtidas para a classe <i>Ritmo Sinusal Normal (RSN)</i> no experimento robusto de classificação multiclasse. . . . .	83
10	Métricas obtidas para a classe <i>Outros</i> no experimento robusto de classificação multiclasse. . . . .	83
11	Métricas média entre todas as classes no experimento robusto de classificação multiclasse. . . . .	84
12	Melhores métricas obtidas no experimento robusto de classificação multilabel. Melhor resultados para cada métrica/classe está destacado em negrito. Métricas sombreadas indicam falta diferença estatística ( $\alpha = 5\%$ ) em relação à melhor métrica. . . . .	85

13	Métricas obtidas para o experimento preliminar de classificação binária. . .	105
14	Métricas obtidas para a classe FA no experimento preliminar de classifica- ção multiclasse. . . . .	106
15	Métricas obtidas para a classe FA no experimento preliminar de classifica- ção multilabel. . . . .	107
16	Métricas obtidas para o experimento robusto de classificação multilabel. . .	108
17	Diagnósticos pontuados no PhysioNet Challenge 2021. . . . .	113

## LISTA DE SIGLAS

<b>FA</b>	Fibrilação Atrial
<b>FLU</b>	Flutter Atrial
<b>AVE</b>	Acidente Vascular Encefálico
<b>ECG</b>	Eletrocardiograma
<b>PTM</b>	Potencial de Transmembrana
<b>SA</b>	sinoatrial
<b>AV</b>	atrioventricular
<b>PA</b>	Potencial de Ação
<b>RSN</b>	Ritmo Sinusal Normal
<b>TS</b>	Taquicardia Sinusal
<b>BS</b>	Bradicardia Sinusal
<b>SNR</b>	Relação Sinal-Ruído

# SUMÁRIO

<b>1</b>	<b>Introdução</b>	<b>16</b>
1.1	Objetivos . . . . .	20
1.2	Organização do documento . . . . .	21
<b>2</b>	<b>Conceitos Fundamentais</b>	<b>22</b>
2.1	Eletrocardiograma . . . . .	22
2.1.1	Eletrofisiologia . . . . .	24
2.1.2	Ritmos de interesse . . . . .	31
2.1.2.1	Ritmo sinusal normal . . . . .	31
2.1.2.2	Taquicardia sinusal . . . . .	32
2.1.2.3	Bradicardia sinusal . . . . .	32
2.1.2.4	Fibrilação atrial . . . . .	33
2.1.2.5	Flutter atrial . . . . .	35
2.1.3	Sistemas de aquisição . . . . .	36
2.1.4	Dispositivos vestíveis . . . . .	39
2.1.4.1	Desafios técnicos . . . . .	41
2.2	Aprendizado de máquina . . . . .	42
2.2.1	Redes Neurais Artificiais . . . . .	44
2.2.2	Deep learning . . . . .	46
2.2.3	Redes Neurais Convolucionais . . . . .	46
<b>3</b>	<b>Metodologia</b>	<b>52</b>
3.1	Base de dados . . . . .	53
3.2	Pré-processamentos . . . . .	56

3.3	Modelos de aprendizado profundo . . . . .	59
3.3.1	Modelos propostos . . . . .	61
3.3.1.1	LiteVGG . . . . .	62
3.3.1.2	LiteResNet . . . . .	66
3.3.1.3	DwNet . . . . .	66
3.4	Métricas de custo computacional . . . . .	67
3.5	Métricas de classificação . . . . .	68
3.6	Experimentos . . . . .	71
3.6.1	Classificação Binária . . . . .	72
3.6.2	Classificação Multiclasse . . . . .	72
3.6.3	Classificação Multilabel . . . . .	73
3.7	Hardwares utilizados . . . . .	74
<b>4</b>	<b>Resultados e discussões</b>	<b>76</b>
4.1	Custo computacional . . . . .	76
4.2	Classificação binária . . . . .	79
4.3	Classificação multiclasse . . . . .	81
4.4	Classificação multilabel . . . . .	84
4.5	Discussões da Classificação . . . . .	86
4.6	Impacto do pré-processamento . . . . .	87
<b>5</b>	<b>Conclusão</b>	<b>89</b>
5.1	Contribuições . . . . .	90
	<b>Referências</b>	<b>91</b>
	<b>Apêndice A – Produção Científica</b>	<b>100</b>
	<b>Apêndice B – ResNet vs LiteResNet</b>	<b>101</b>

<b>Apêndice C – Resultados dos experimentos preliminares</b>	<b>104</b>
<b>Apêndice D – Resultados do experimento multilabel robusto</b>	<b>108</b>
<b>Apêndice E – Trabalhos futuros</b>	<b>112</b>
<b>Anexo A – Rótulos do dataset PhysioNet Challenge 2021</b>	<b>113</b>



# 1 INTRODUÇÃO

Segundo o estudo GLOBAL BURDEN OF DISEASE COLLABORATIVE NETWORK (2020), doenças cardiovasculares são a principal causa de morte ao redor do mundo, sendo responsáveis por 32,84% das mortes. No Brasil, as doenças cardiovasculares são responsáveis por 28,21% das mortes. Dentre as doenças cardiovasculares, a doença isquêmica do coração e o Acidente Vascular Encefálico (AVE) são responsáveis pela maioria das mortes.

Tabela 1: Percentual de mortes relacionadas a doenças cardiovasculares.

	<b>Causa</b>	<b>Mundo (%)</b>	<b>Brasil (%)</b>
	Doença isquêmica do coração	16,17	12,14
	Acidente Vascular Encefálico	11,59	9,28
	Doença cardíaca hipertensiva	2,05	2,11
	Cardiomiopatia e miocardite	0,60	1,52
	Fibrilação e flutter atrial	0,56	0,77
	Doença cardíaca reumática	0,54	0,19
	Aneurisma da aorta	0,31	0,75
	Doença na artéria periférica	0,13	0,20
	Endocardite	0,12	0,17
	Valvulopatia não reumática	0,29	0,34
	Outras doenças cardiovasculares e circulatórias	0,49	0,74
	<b>Doenças cardiovasculares (TOTAL)</b>	<b>32,84</b>	<b>28,21</b>

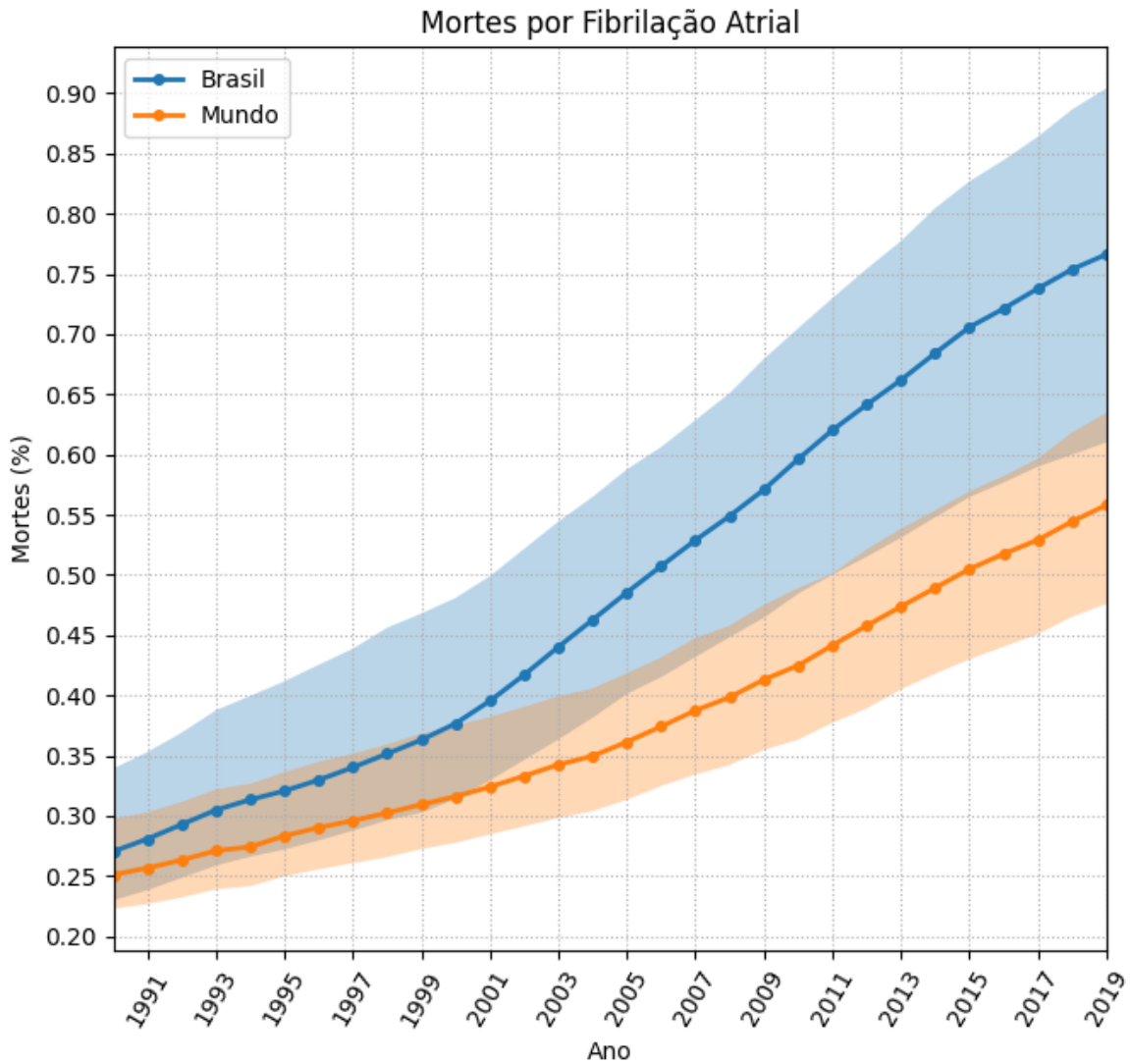
Fonte: GLOBAL BURDEN OF DISEASE COLLABORATIVE NETWORK (2020).

Dentre as mortes relacionadas às doenças cardiovasculares, a Fibrilação Atrial e o Flutter Atrial, ambas abreviadas por FA, são responsáveis por 0,77% e 0,56% das mortes no Brasil e no mundo, respectivamente. Entretanto, a FA é um fator de risco associado com o aumento a longo prazo na probabilidade de insuficiência cardíaca e AVE (ISSA et al., 2009; CRANDALL et al., 2009). O diagnóstico de FA também é associado a um risco de morte 1,5 a 1,9 vezes maior do que em pacientes sem arritmia, em parte devido a uma forte associação a eventos tromboembólicos (BABAR, 2022).

Apesar do baixo número de mortes associadas a FA, vale destacar que este número

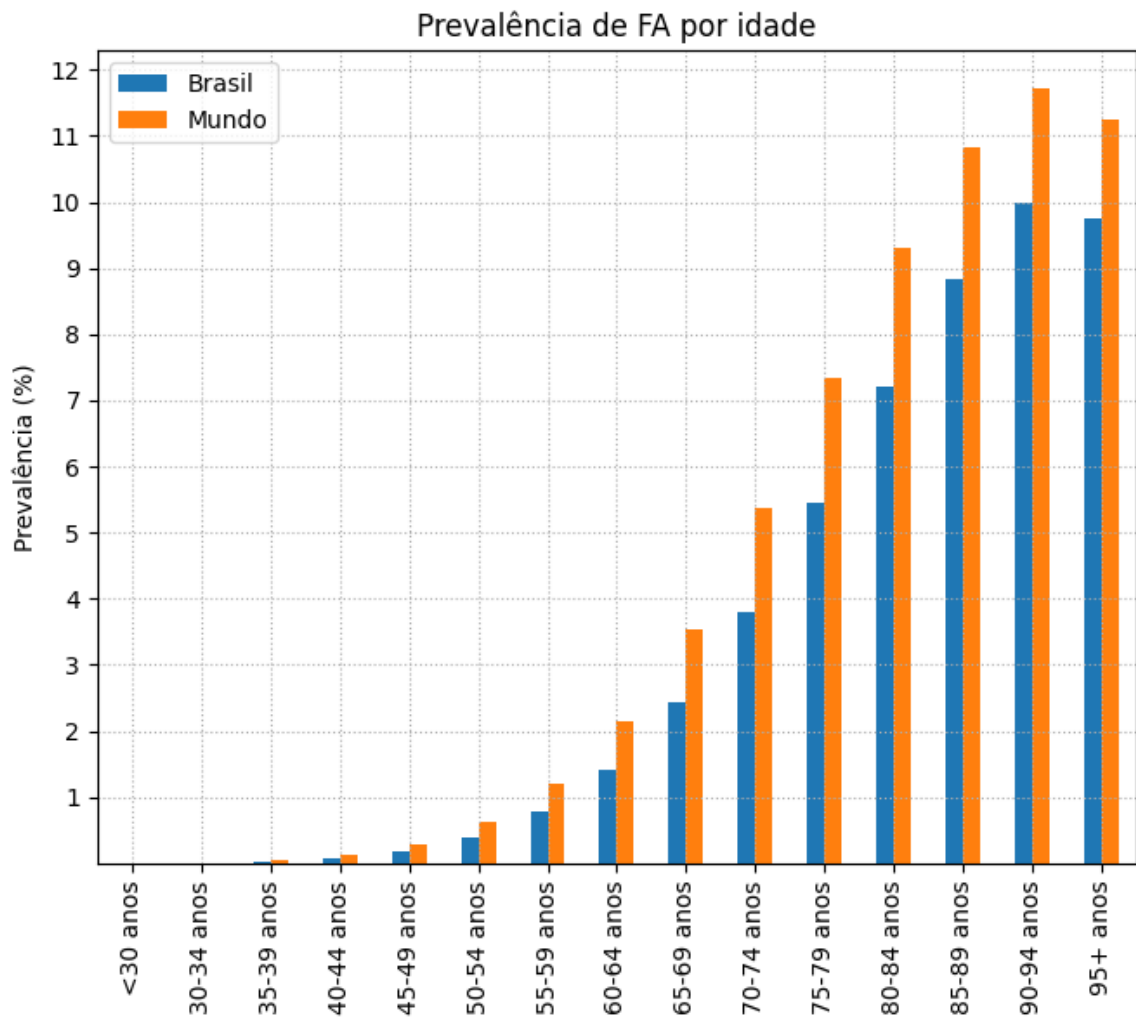
vem crescendo nos últimos 30 anos, conforme podemos observar na Figura 1. A tendência da prevalência de FA é de um aumento conforme a média de idade da população aumenta (KORNEJ et al., 2020). A Figura 2 exhibe a relação da prevalência por faixa etária no Brasil e no mundo. A população mais idosa naturalmente possui uma prevalência elevada, sendo que ao redor do mundo esta prevalência é superior a 5% para pessoas acima de 70 anos e para pessoas acima de 85 anos esta prevalência mais do que dobra (GLOBAL BURDEN OF DISEASE COLLABORATIVE NETWORK, 2020). Vale destacar que as estimativas de prevalência podem estar subestimadas, visto que cerca de 10 a 15% dos casos são assintomáticos (MAGALHÃES et al., 2016).

Figura 1: Percentual de mortes causadas por Fibrilação Atrial ao longo dos anos. Área sombreada indica intervalo de confiança.



Fonte: GLOBAL BURDEN OF DISEASE COLLABORATIVE NETWORK (2020).

Figura 2: Prevalência de Fibrilação Atrial em 2019, por grupos de idade.



Fonte: GLOBAL BURDEN OF DISEASE COLLABORATIVE NETWORK (2020).

Este aumento na prevalência ocorre pois a prevalência de FA já é maior para a população idosa, sendo superior a 5% para pessoas acima de 75 anos e atingindo mais de 10% para pessoas acima de 85 anos. Esta relação da prevalência por faixa etária é exibida na Figura 2. Outros fatores que colaboram para esta tendência de incremento na prevalência de FA são os avanços no tratamento de doenças cardiovasculares crônicas, o que aumenta o tamanho da população mais suscetível a FA, e a maior disponibilidade de exames para monitoramento prolongado do ritmo cardíaco, como o Holter, aumentando a detecção de casos assintomáticos (MAGALHÃES et al., 2016).

O diagnóstico precoce da FA é um fator chave para um correto tratamento, em especial para pacientes que já possuem alguma outra cardiopatia ou fatores de alto risco, como

hipertensão e diabetes (AHMED; ZHU, 2020; WORLD HEALTH ORGANIZATION, 2021). O monitoramento contínuo do Eletrocardiograma (ECG) por meio de dispositivos vestíveis, que permitem o paciente continuar naturalmente com suas atividades cotidianas, pode ser uma das alternativas para obter este diagnóstico precoce (JOUNG et al., 2019).

O diagnóstico rápido de eventos de FA é extremamente importante para pacientes após cirurgias cardíacas. Nesta população, a incidência varia de 10% a 65%, dependendo do tipo de cirurgia realizada (PASSANNANTE, 2011; DAVE et al., 2018). O pico de incidência está entre dois e quatro dias depois da cirurgia (CAMANHO; SANTOS, 2019), porém alguns pacientes podem apresentar FA depois de seis ou mais dias após a cirurgia (ARANKI et al., 1996).

O gerenciamento da FA no pós-operatório cardíaco tem impacto no período de internação dos pacientes, e conseqüentemente no gasto de recursos hospitalares. O diagnóstico precoce da FA pode auxiliar nos esforços para reduzir o tempo de internação dos pacientes, resultando em redução de custos e diminuindo a morbidade do paciente (PEREIRA et al., 2020; ARANKI et al., 1996).

Uma das dificuldades no monitoramento contínuo do ECG é o grande volume de dados a ser analisado pelo cardiologista. Em geral, cardiologistas utilizam ferramentas computacionais de análise para auxiliarem nos diagnósticos de exames de longa duração. Tais ferramentas identificam áreas com possíveis eventos, sendo que o cardiologista é responsável por realizar a correta identificação do evento. Atualmente existem algumas ferramentas de telemonitoramento assistido por computador que envolvem o envio de todos os sinais coletados para um servidor, onde modelos computacionais robustos são capazes de gerar alertas em casos de detecção de arritmias (MEZOO, 2022; PROLIFE, 2022; VIVALINK, 2022).

Nestas ferramentas de telemonitoramento, o envio dos sinais para análise do ritmo cardíaco geralmente é realizado por meio de um dispositivo vestível conectado à Internet. Desse modo, é necessária uma conexão estável de Internet para que o monitoramento em tempo real do ritmo possa gerar alertas para o paciente, o que nem sempre é a realidade, especialmente para pacientes de baixa renda. Neste modelo de funcionamento existe a possibilidade, ainda pouco explorada, de utilizar dispositivos móveis do tipo smartphone para realizar a análise do ritmo cardíaco. Esta alternativa diminui os custos operacionais de um sistema de telemonitoramento do ritmo cardíaco, visto que os sinais de ECG podem ser transmitidos apenas quando a análise neste tipo de dispositivo indicar possibilidade

de arritmia. Além disso, o sistema de alerta estaria disponível para o paciente mesmo sem uma conexão de Internet.

Uma dificuldade encontrada na integração de um sistema de análise de ritmos cardíacos inteiramente dentro de uma aplicação móvel está no tamanho e complexidade dos modelos usualmente utilizados nestas análises, em especial aqueles baseados em aprendizado profundo, que se destacam com melhores acurácias do que modelos de aprendizado de máquina clássicos (REYNA et al., 2022) e não necessitam de etapas de extração de características. Modelos cada vez mais complexos são utilizados para melhorar a acurácia na identificação de arritmias (SUH et al., 2021; MAGNI et al., 2021), sendo que uma das opções com melhor acurácia é a utilização de mais de um modelo e utilizar uma estratégia de votação para decidir a saída final da classificação (NEJEDLY et al., 2021; REYNA et al., 2022). Em todos estes casos, os modelos são criados sem levar em consideração custos computacionais e outras limitações de plataformas móveis (KATSAOUNI et al., 2022), visto que atualmente as tecnologias de telemonitoramento estão voltadas para soluções em nuvem, onde podem ser utilizados servidores com maior poder computacional.

O desenvolvimento de modelos de baixo custo computacional para a classificação de ritmo cardíaco possibilitaria explorar uma arquitetura de monitoramento cardíaco móvel menos dependente de conexão com internet, fornecendo alertas em tempo-real e gerando um menor custo tanto em termos de pacotes de dados por parte dos pacientes quanto em termos de infraestrutura de rede por parte de hospitais e provedores de serviços de telecardiologia.

## 1.1 Objetivos

Este trabalho tem como objetivo principal o desenvolvimento de um modelo classificador de sinais de ECG de baixo custo computacional para identificação de ritmos cardíacos do tipo FA.

Para tanto, são exploradas algumas arquiteturas de aprendizado profundo (*deep learning*) já consolidados e propostas alterações para redução no custo computacional dos mesmos. Embora o principal alvo da análise de ritmos seja a FA, outros ritmos (sinusal normal, bradicardia e taquicardia) também são considerados na análise.

Um pré-processamento robusto de sinais de ECG também fará parte deste classificador, levando em consideração uma possível aplicação deste classificador em sinais com qualidade reduzida, como aqueles adquiridos por ECG vestível com o paciente executando

suas atividades diárias.

## 1.2 Organização do documento

A dissertação foi organizada em cinco capítulos apresentados conforme a descrição a seguir:

O Capítulo 1, Introdução, contém o tema da pesquisa, suas motivações, delimitações e objetivos, ou seja, a presente dissertação objetiva desenvolver um algoritmo de baixo custo computacional para classificação de ritmos cardíacos, em especial FA, a partir de sinais de ECG vestíveis.

O Capítulo 2, Conceitos Fundamentais, contém uma revisão dos conceitos necessários para a compreensão desta dissertação. Estes conceitos estão relacionados à eletrocardiografia, incluindo princípios de eletrofisiologia, sistemas de aquisição e ritmos de interesse; e relacionados ao aprendizado de máquina, em especial princípios de aprendizado profundo.

O Capítulo 3, Metodologia, discorre sobre a abordagem utilizada na pesquisa relatada nesta dissertação, incluindo bases de dados utilizadas, métricas de classificação, pré-processamentos aplicados, arquitetura de modelos de aprendizado profundo e estratégias de treinamento para classificação.

No Capítulo 4, Resultados e discussões, são analisados os resultados obtidos para os modelos propostos e é realizada uma comparação com modelos clássicos, mais complexos. Nestas análises são consideradas o custo computacional, os resultados da classificação e o impacto do pré-processamento utilizado.

Por fim, o Capítulo 5, Conclusão, apresenta as considerações finais desta pesquisa, listando também as contribuições obtidas.

## 2 CONCEITOS FUNDAMENTAIS

Este capítulo apresenta uma introdução a conceitos fundamentais que serão abordados ao longo do texto. Estes conceitos estão agrupados em dois temas principais: eletrocardiografia e aprendizado de máquina.

Referente ao tema de eletrocardiografia, serão abordados: eletrofisiologia, conceitos de sistemas de aquisição, ritmos de interesse para este trabalho e a aquisição por meio de dispositivos vestíveis.

Em relação a aprendizado de máquina, serão abordados: redes neurais artificiais e aprendizado profundo.

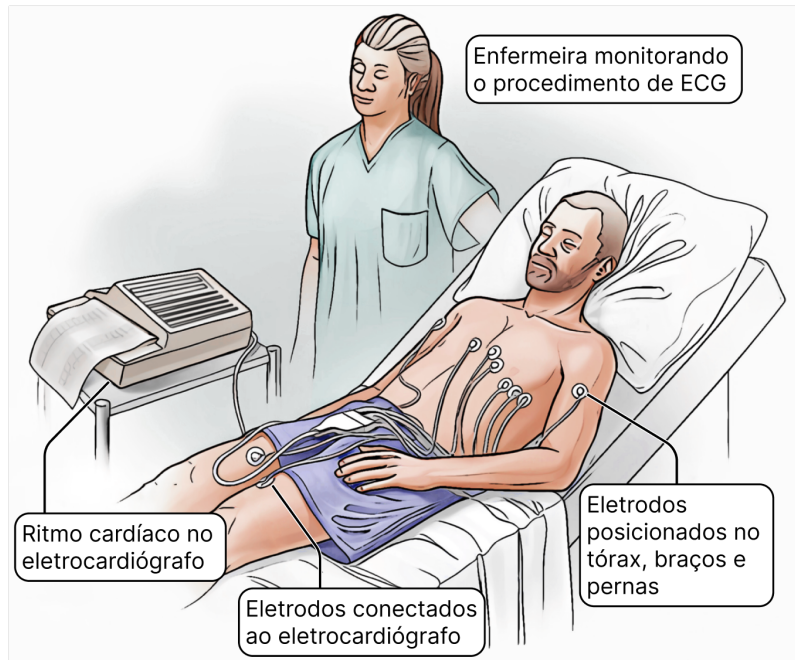
### 2.1 Eletrocardiograma

De forma sucinta, o ECG é um exame que faz o registro temporal da diferença de potencial elétrico causada pela atividade cardíaca e que é adquirida de forma não invasiva na superfície do corpo. No contexto clínico, o ECG usualmente é adquirido como 12 diferenças de potenciais, ou derivações.

A Figura 3 ilustra como é realizado o procedimento para a aquisição dos sinais de ECG de repouso. O paciente é acomodado em uma maca e eletrodos são acoplados nos braços, pernas e tórax. Durante o procedimento, o paciente deve evitar qualquer movimento e um técnico monitora todo o procedimento para verificar possíveis desconexões de eletrodos ou movimentações do paciente. A aquisição dos sinais geralmente é realizada durante um intervalo de 10 segundos. Eletrocardiógrafos antigos registravam estes sinais diretamente em papel, mas equipamentos modernos registram digitalmente o sinal para depois imprimir os traçados, possibilitando assim processamentos dos sinais adquiridos (KLIGFIELD et al., 2007; SATTAR; CHHABRA, 2023). A Figura 4 mostra a forma de onda de um ECG típico, exibindo as suas principais componentes (ondas P, Q, R, S e T) e intervalos de interesse. A Figura 5 mostra um exemplo real de ECG de 12 derivações

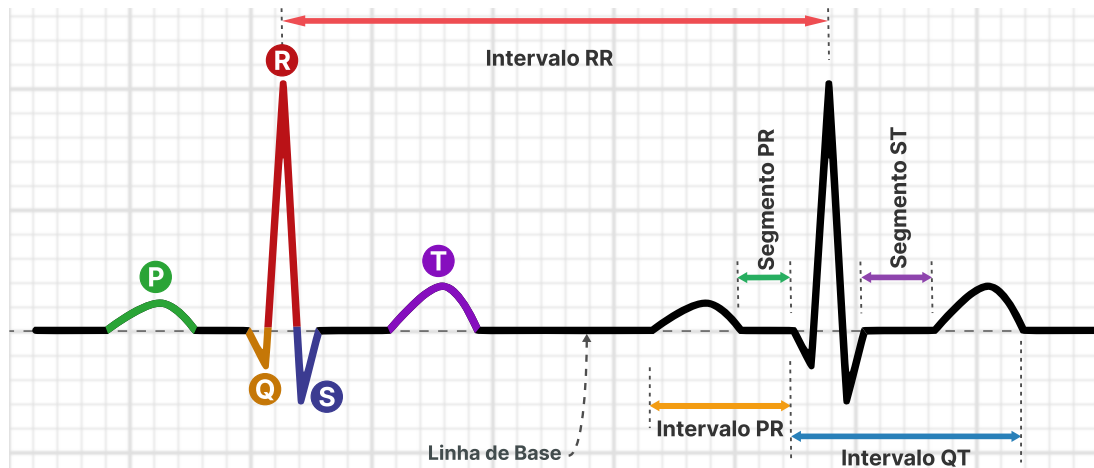
de um paciente com ritmo normal, observa-se que a forma de onda típica é válida apenas para a derivação II e um ECG real eventualmente ainda possuirá ruídos e artefatos.

Figura 3: Procedimento de aquisição de ECG de repouso.



Fonte: Adaptado de BANGKOK DUSIT MEDICAL SERVICES (2022).

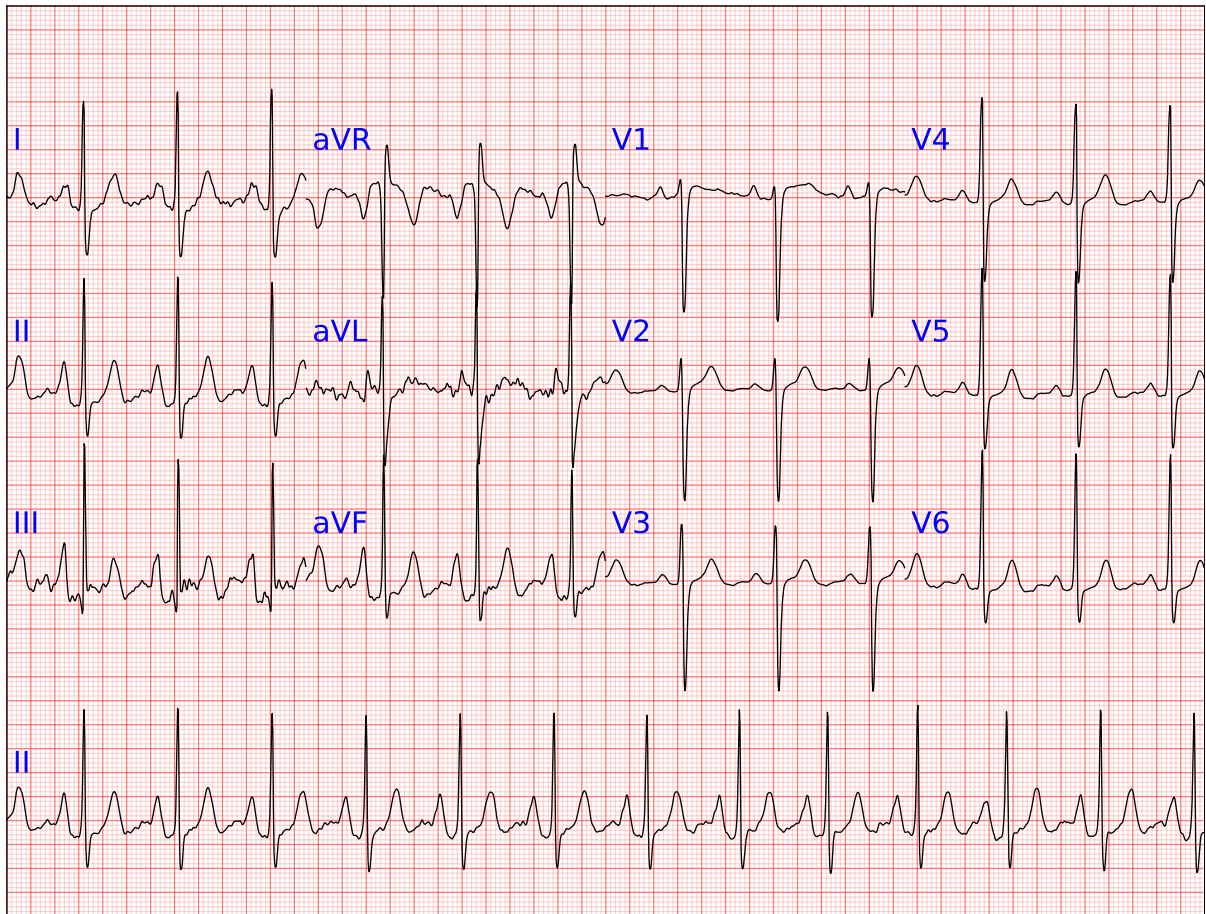
Figura 4: Formato de onda de um ECG típico.



Fonte: Adaptado de Atkielski (2022) e Guyton e Hall (2011).



Figura 5: Exemplo de um ECG com ritmo normal.



Fonte: Amostra obtida do banco de dados publicado por Reyna et al. (2021) (ID A0002).

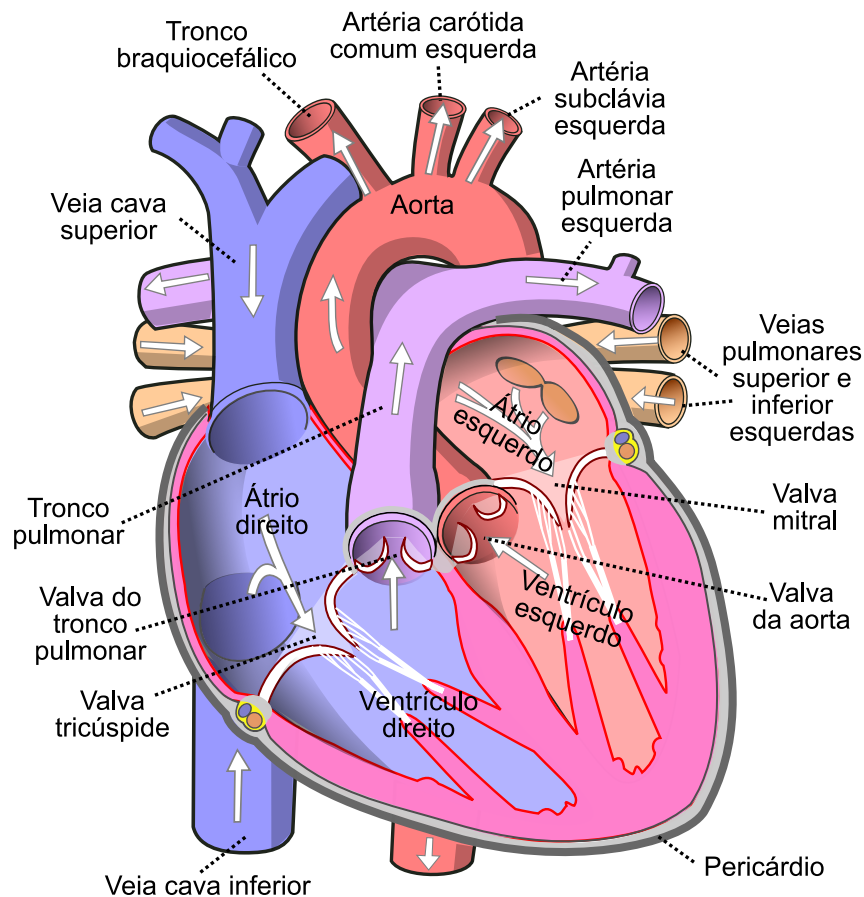
### 2.1.1 Eletrofisiologia

O sistema cardiovascular é responsável pelo transporte de substâncias dentro do corpo, levando sangue oxigenado do pulmão para o restante do corpo, removendo gás carbônico e outros resíduos dos corpos e transportando nutrientes e hormônios (LOPES, 2004). O coração funciona como uma bomba hidráulica, fornecendo o trabalho necessário para o sangue se deslocar nas artérias e veias (WEBSTER, 1965; FURST, 2015).

O sangue com menor concentração de oxigênio chega ao átrio direito através das veias cavas superior e inferior. O sangue preenche o átrio direito enquanto o mesmo está relaxado (fase de diástole) e preenche parcialmente o ventrículo direito, sendo que o ventrículo é completamente preenchido durante a sístole do átrio. Após a contração do átrio direito, inicia-se a contração do ventrículo direito, durante esta fase a valva tricúspide se fecha e a pressão do sangue no interior do ventrículo aumenta até atingir uma pressão suficiente para abrir a valva do tronco pulmonar, expulsando o sangue para

o pulmão através do tronco pulmonar. O sangue então passa pelo pulmão, onde ocorrem as trocas gasosas e o sangue se torna mais oxigenado. O sangue mais oxigenado é coletado pelas veias pulmonares e preenchem o átrio esquerdo. De forma semelhante ao que ocorre no lado direito, o ventrículo esquerdo é parcialmente preenchido pelo fluxo e pressão residual do sangue venoso, mas só é completamente preenchido após a contração do átrio esquerdo. Após a contração do átrio esquerdo, ocorre a contração do ventrículo esquerdo, aumentando a pressão sanguínea dentro do ventrículo, fechando a valva mitral e abrindo a valva da aorta após a pressão do ventrículo atingir uma pressão superior à pressão sanguínea da aorta. O sangue é então expelido pela aorta, distribuído pelo corpo e é finalmente coletado pelas veias até as veias cavas superior e inferior, completando o ciclo do sangue no corpo (GUYTON; HALL, 2011; SATTAR; CHHABRA, 2023). A Figura 6 mostra um diagrama do coração humano, exibindo todas as estruturas citadas neste parágrafo e com setas brancas indicando a direção do sangue em cada uma das veias, artérias e câmaras do coração.

Figura 6: Diagrama do coração humano.



Fonte: Nogué (2022).

Para o correto funcionamento do sistema circulatório é necessário uma coordenação dos momentos de contração e relaxamento dos átrios e ventrículos do coração. A contração e relaxamento dos átrios direito e esquerdo ocorrem praticamente de forma simultânea, sendo que a contração dos ventrículos deve ocorrer após a contração dos átrios. O atraso entre a contração atrial e ventricular deve ser suficiente para o sangue ser expulso dos átrios para os ventrículos. A descoordenação entre as contrações pode resultar em insuficiência cardíaca (devido à redução no volume sistólico), formação de trombos, ou outras complicações a depender do tipo de descoordenação.

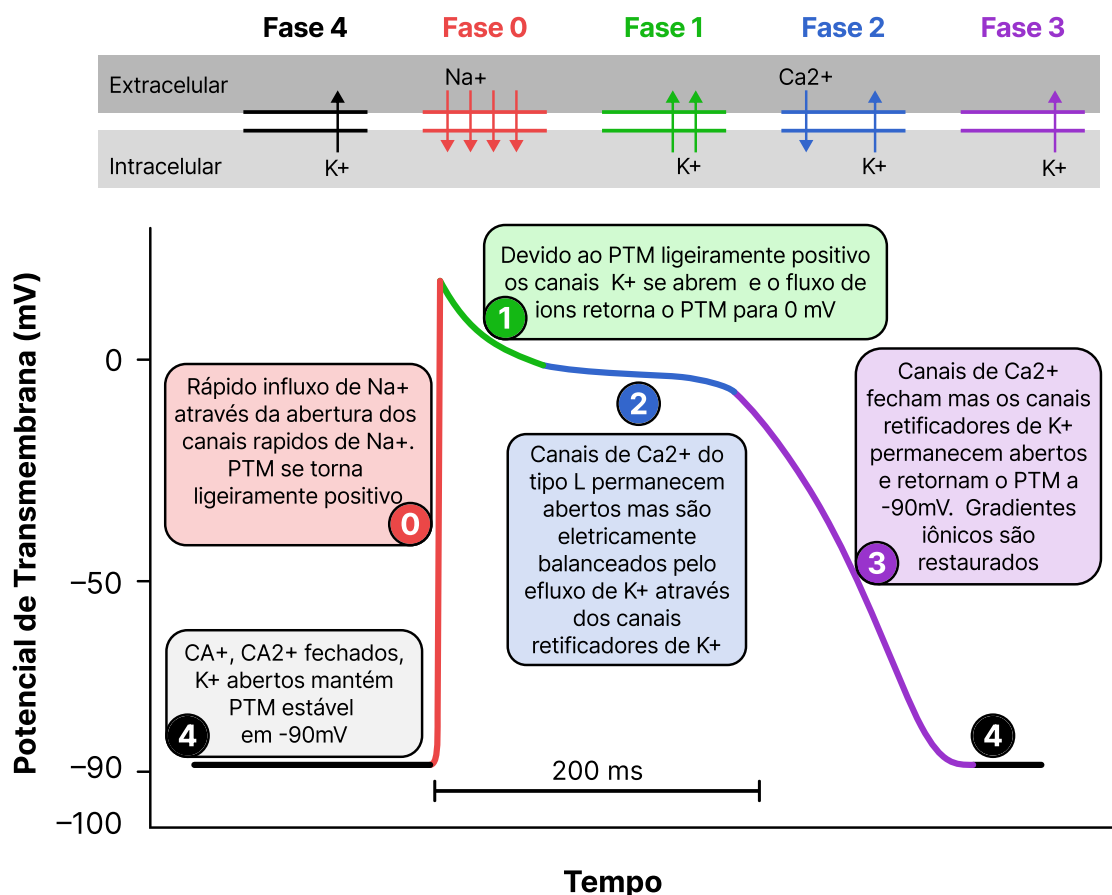
A coordenação, e também a periodicidade, das contrações das câmaras do coração é controlada por impulsos elétricos que se propagam como variações no Potencial de Transmembrana (PTM). O PTM é a diferença de potencial elétrico entre o meio intra e extra celular, causado por diferenças na concentração de íons  $\text{Na}^+$ ,  $\text{Ca}^{2+}$  e  $\text{K}^+$  em cada um dos meios. O fluxo de íons é governado pelo potencial elétrico, íons de mesma carga se repelem, e pelo potencial químico, íons tendem a se movimentar de forma a reduzir o gradiente de concentração. De forma passiva, a tendência é que o PTM seja zero, entretanto, canais de íons presentes na membrana das células cardíacas geram fluxos de íons que mantêm o gradiente de concentração iônica e, conseqüentemente, uma diferença de potencial elétrico. Cada tipo de canal de íon é permeável a íons específicos, sendo que a abertura e fechamento ocorre após certos limiares de tensão do PTM e alguns ainda possuem uma dependência temporal (apresentando um período refratário entre aberturas consecutivas). Dessa forma, diferentes canais de íons influenciam em diferentes níveis cada uma das fases da propagação do impulso elétrico através das células cardíacas. (AMIN et al., 2010)

O impulso elétrico que se propaga por meio das variações do PTM também é conhecido como Potencial de Ação (PA). A Figura 7 mostra as cinco fases do PA de um cardiomiócito. A seguir são descritas cada uma destas fases (NERBONNE; KASS, 2005):

- Fase 0 - Despolarização: O PA se inicia devido a um valor de PTM superior a  $-90$  mV em uma célula vizinha. Neste momento canais rápidos de  $\text{Na}^+$  se abrem aumentando o PTM do cardiomiócito, sendo que ao atingir  $-70$  mV ocorre um fluxo autossustentado de íons  $\text{Na}^+$ . Esta rápida corrente iônica resulta na despolarização e o PTM chega a atingir um potencial positivo. Após a despolarização os canais rápidos de  $\text{Na}^+$  se fecham e iniciam seu período refratário, durante o qual não se abrirá mesmo com um PTM superior a  $-90$  mV. Canais de  $\text{Ca}^{2+}$  se abrem quando o PTM ultrapassa  $-40$  mV.

- Fase 1 - Repolarização inicial: Com o PTM ligeiramente positivo, alguns canais de  $K^+$  se abrem e o fluxo de íons retorna o PTM para 0 mV.
- Fase 2 - Plateau: Os canais de  $Ca^{2+}$  continuam abertos e bombeiam íons positivos para o meio intracelular, mas canais de  $K^+$  permitem o fluxo de íons positivos para o meio extracelular. Desse modo, os fluxos iônicos se balanceiam eletricamente e mantém o PTM próximo a 0 mV.
- Fase 3 - Repolarização: Os canais de  $Ca^{2+}$  começam a fechar, mas os canais de  $K^+$  continuam bombeando íons positivos para o meio extracelular, retornando assim o PTM para o potencial de repouso de -90 mV. O gradiente de concentração iônica de  $Ca^{2+}$  e  $Na^+$  é restaurado para o meio extracelular, enquanto o gradiente de  $K^+$  é restaurado para o meio intracelular.
- Fase 4 - Repouso: Durante esta fase, os canais de  $Ca^{2+}$  e  $Na^+$  se mantêm fechados, enquanto o PTM se mantém constante por meio dos canais de  $K^+$ .

Figura 7: Fases do potencial de ação de células cardíacas.



Fonte: Adaptado de Ikonnikov e Wong (2013) e Nattel (2003).

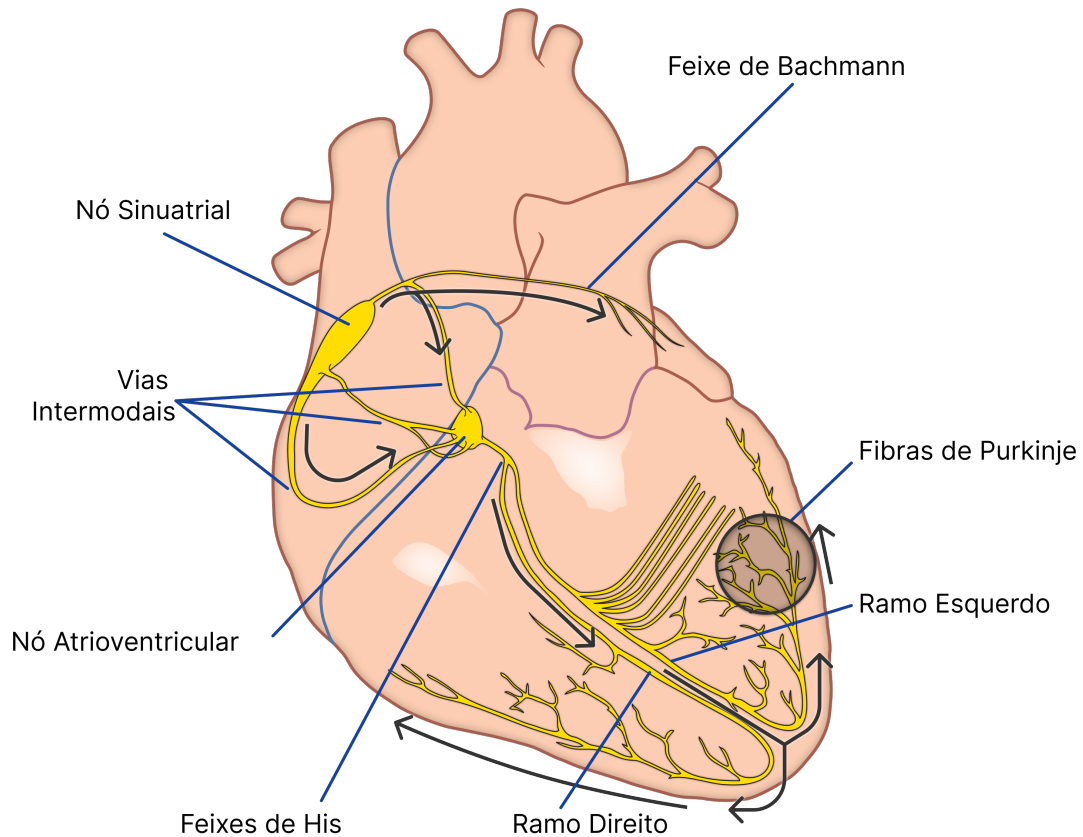
O comportamento do PA descrito acima necessita de uma excitação inicial. Esta excitação normalmente é feita por meio de células marca-passo, as quais são autoexcitáveis. Células marca-passo estão presentes no nó SA, nó atrioventricular (AV) e ventrículos. Entretanto, cada local possui uma autoexcitação com frequências diferentes. As células do nó SA são o marca-passo principal do coração e, por possuir maior frequência de disparo, governam o ritmo cardíaco em condições normais (DIFRANCESCO; BORER, 2007).

Considerando condições normais, o PA gerado espontaneamente no nó SA é propagado através do átrio direito e, por meio do feixe de Bachmann, até o átrio esquerdo, estimulando a contração muscular dos átrios. Ramos especiais, chamados vias intermodais, são responsáveis por propagar este PA até o nó AV, onde ocorre um atraso na propagação do PA. Este atraso é fundamental para coordenar as contrações atriais e ventriculares, de modo que o sangue seja ejetado dos átrios para os ventrículos e a contração ventricular ocorra posteriormente. Outra função importante do nó AV é sua condução unidirecional do PA, o que impede, em condições normais, o retorno de um estímulo elétrico dos ventrículos para os átrios. O PA então é propagado do nó AV para os feixes de His, que se ramificam nos ramos direito e esquerdo. As fibras de Purkinje presentes nos ramos direito e esquerdo são responsáveis por propagar rapidamente o PA através dos ventrículos e realizar a estimulação dos cardiomiócitos do endocárdio. O PA então é propagado até as superfícies epicárdicas com uma velocidade menor. Devido à rápida propagação do PA no endocárdio dos ventrículos, as fibras musculares se contraem de forma praticamente sincronizada no ventrículo como um todo. Como todos os cardiomiócitos possuem a capacidade de condução de PA, o ânulo fibroso (ou esqueleto cardíaco) funciona como uma isolação elétrica entre os músculos de átrios e ventrículos (ANDERSON et al., 2009; GUYTON; HALL, 2011; CARMO, 2022).

A Figura 8 mostra as estruturas relacionadas ao sistema de condução de PA apresentadas no parágrafo anterior, sendo que o sentido de propagação do PA é denotado por setas pretas ao longo das estruturas. A Figura 9 mostra o tempo necessário para o impulso gerado no nó SA se propagar a diferentes regiões do coração, nesta figura é possível observar como o nó AV gera um atraso de aproximadamente 0,1 segundo entre a despolarização dos átrios e dos ventrículos.

O PA se propaga ao longo do coração, entretanto, a superposição dos potenciais individuais de cada célula se propaga até a superfície do corpo, sofrendo atenuações diversas de acordo com as impedâncias dos tecidos pelos quais o campo elétrico está sendo propagado. O potencial elétrico mensurado através de técnicas de ECG apresenta uma correlação com os PA de cada região do coração, conforme é mostrado na Figura 10. A

Figura 8: Sistema de condução de potenciais de ação do coração.

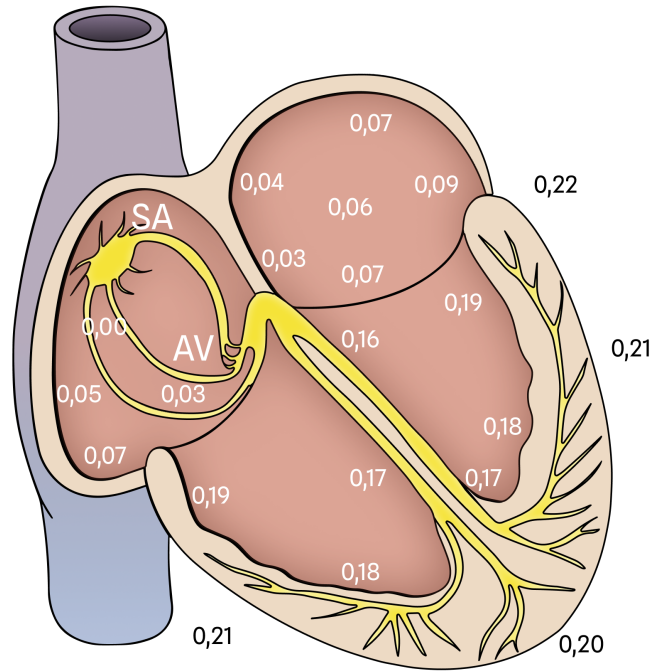


Fonte: Adaptado de Madhero88 (2010).

chamada onda P do ECG está relacionada com a despolarização dos átrios, o complexo QRS com a despolarização dos ventrículos e a onda T com a repolarização dos ventrículos (VARRÓ et al., 2021; GUYTON; HALL, 2011). Entretanto, a morfologia do sinal de ECG também varia de acordo com os dipolos utilizados para a aquisição dos sinais, com as características anatômicas do paciente, com deslocamentos gerados pela respiração e outros artefatos de movimentação.

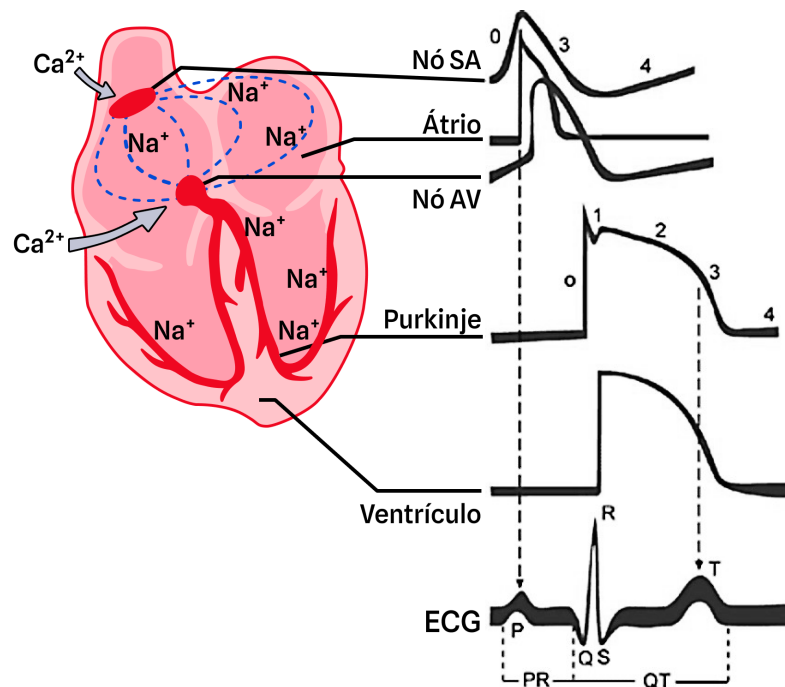
A Figura 11 exibe a correlação do ECG com eventos do ciclo cardíaco. Como a atividade elétrica é responsável pela coordenação do ciclo cardíaco, alterações no mesmo são indícios de possíveis cardiopatias. No entanto, nem toda cardiopatia gera alterações no ECG, sendo necessários outros exames para suas detecções. Dentre as cardiopatias detectáveis por ECG, destaco nesta dissertação aquelas de ordem rítmica e de condução elétrica, como arritmias, bigeminia e bloqueios.

Figura 9: Instantes aproximados do início da despolarização dos cardiomiócitos em diferentes partes do coração (em segundos após a despolarização inicial do nó SA).



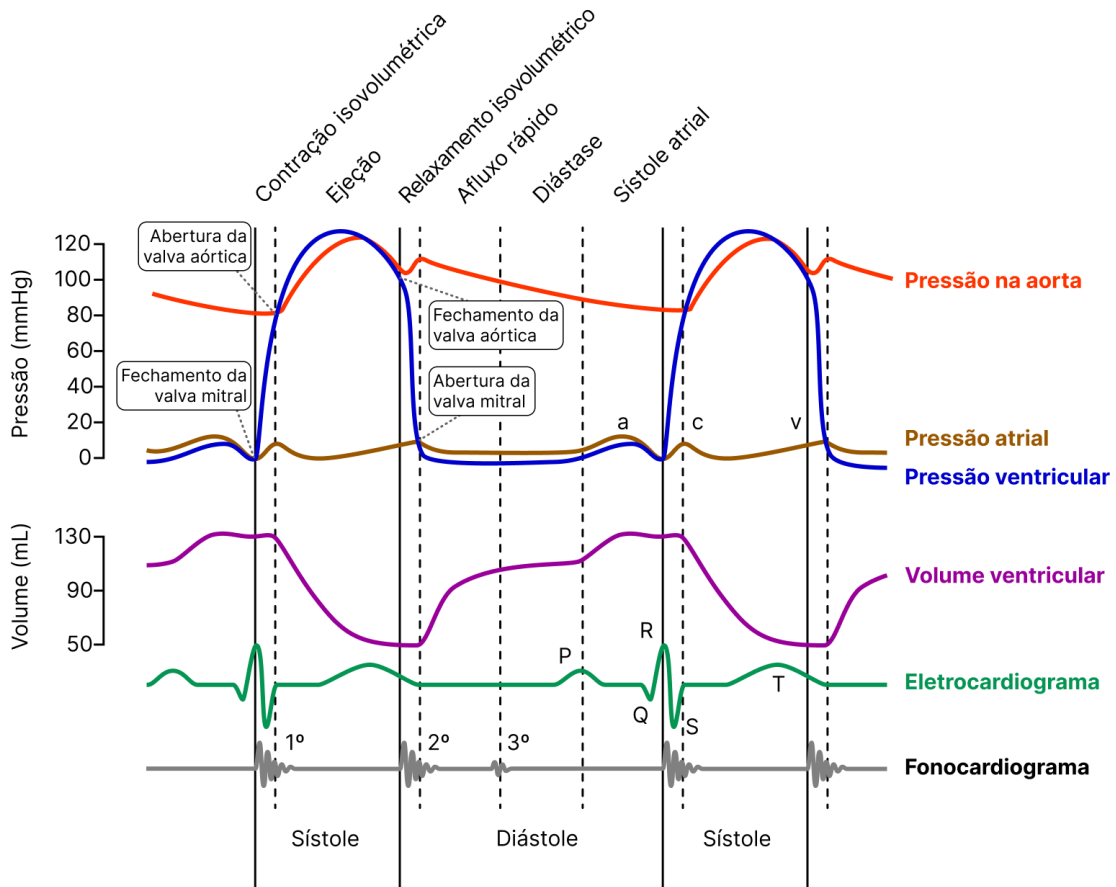
Fonte: Adaptado de Guyton e Hall (2011).

Figura 10: Representação esquemática dos potenciais de ação em diferentes regiões do coração e suas correlações com o ECG.



Fonte: Adaptado de PROYECTO ITACA.

Figura 11: Eventos do ciclo cardíaco, mostrando pressão no ventrículo esquerdo, pressão no átrio esquerdo, pressão na aorta, volume do ventrículo esquerdo, ECG e fonocardiograma.



Fonte: Adaptado de Adh30 et al. (2016) e Guyton e Hall (2011).

## 2.1.2 Ritmos de interesse

Este trabalho considerou a classificação de cinco ritmos cardíacos, sendo o principal destes a FA. Estes ritmos foram escolhidos de forma a possibilitar o acompanhamento de pacientes saudáveis ou com alguma cardiopatia conhecida, possibilitando o disparo de alertas de baixo ou alto risco. A escolha destes ritmos também foi influenciada pela prevalência dos mesmos em bases públicas de ECG anotadas.

### 2.1.2.1 Ritmo sinusal normal

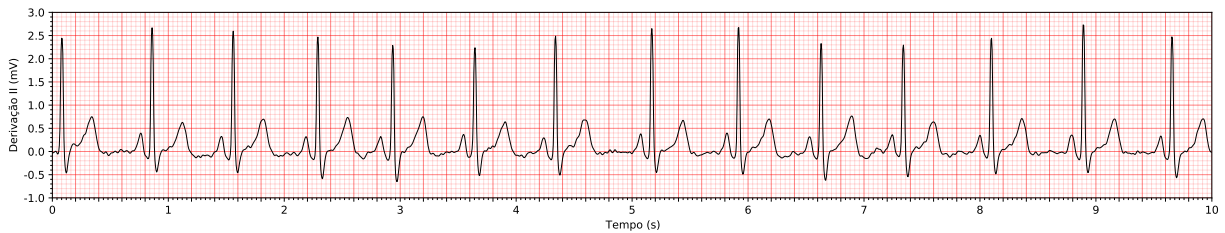
O Ritmo Sinusal Normal (RSN) é o ritmo de um coração saudável em condições normais, em que o impulso elétrico é iniciado no nó SA e propagado de forma adequada para o restante do coração. Tal ritmo é caracterizado por uma onda P com polaridades corretas (positiva na derivação II e negativa na derivação aVR), onda P precedendo cada



complexo QRS e frequência cardíaca entre 60 e 100 bpm (GOLDBERGER et al., 2018).

A Figura 12 exibe a derivação II do ECG de um paciente apresentando ritmo RSN. Observa-se ondas P positivas e sempre precedidas de um complexo QRS, frequência cardíaca de aproximadamente 62 bpm e uma regularidade dos intervalos R-R.

Figura 12: Derivação II do ECG de um paciente com ritmo RSN.



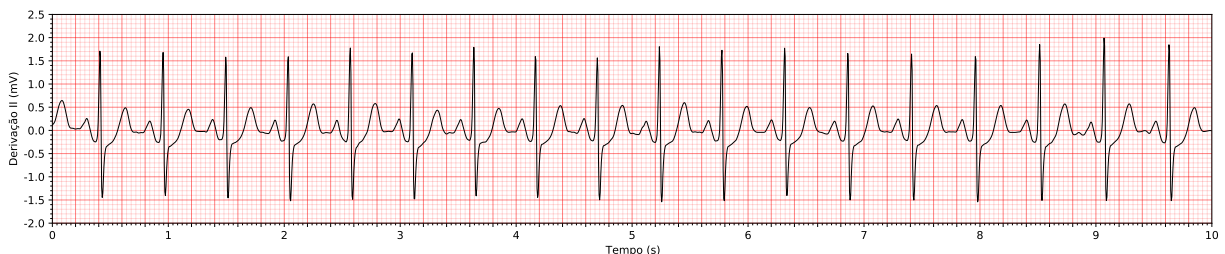
Fonte: Amostra obtida do banco de dados publicado por Reyna et al. (2021) (ID A0073).

### 2.1.2.2 Taquicardia sinusal

A Taquicardia Sinusal (TS) corresponde a um ritmo de origem sinusal com alta frequência cardíaca, superior a 100 bpm (GUYTON; HALL, 2011). Este é um ritmo que pode ocorrer naturalmente e de forma benigna em condições de exercício intenso ou choques emocionais, mas também pode ocorrer devido a febre, perda de volume intravascular, insuficiência cardíaca, entre outras condições (GOLDBERGER et al., 2018).

A Figura 13 exibe a derivação II do ECG de um paciente apresentando ritmo TS. Observa-se uma onda com morfologia muito similar a um ritmo RSN, porém com uma alta frequência cardíaca, aproximadamente 117 bpm.

Figura 13: Derivação II do ECG de um paciente com ritmo TS.



Fonte: Amostra obtida do banco de dados publicado por Reyna et al. (2021) (ID E00011).

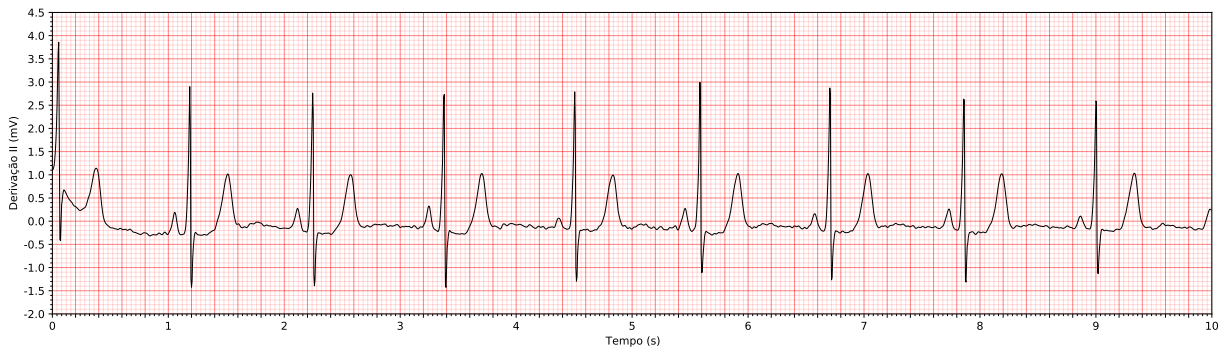
### 2.1.2.3 Bradicardia sinusal

A Bradicardia Sinusal (BS) também é outro ritmo em que o impulso elétrico no nó SA. Entretanto, ao contrário da taquicardia, apresenta uma baixa frequência cardíaca,

inferior a 60 bpm. Este é o ritmo para a condição de repouso de algumas pessoas saudáveis, sendo um ritmo rotineiro durante o sono (GOLDBERGER et al., 2018). A BS também pode estar associada com apneia obstrutiva do sono, uso de medicamentos e anomalias endócrinas e metabólicas (HAFEEZ; GROSSMAN, 2022).

A Figura 14 exibe a derivação II do ECG de um paciente apresentando ritmo BS. Observa-se uma onda com morfologia muito similar a um ritmo RSN, porém apresentando uma baixa frequência cardíaca, aproximadamente 51 bpm.

Figura 14: Derivação II do ECG de um paciente com ritmo BS.



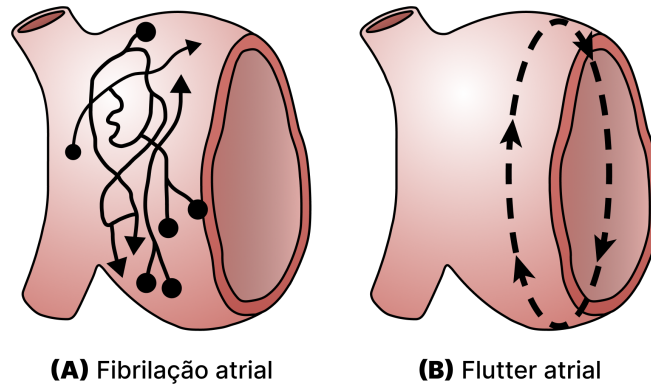
Fonte: Amostra obtida do banco de dados publicado por Reyna et al. (2021) (ID E00536).

#### 2.1.2.4 Fibrilação atrial

A FA é um ritmo no qual o estímulo elétrico gerado pelo nó SA se propaga de forma caótica nos átrios, consequentemente causando uma contração não coordenada da musculatura atrial (NATTEL, 2003). A Figura 15(A) exemplifica como ocorre esta propagação caótica. Neste regime, os impulsos elétricos atingem aleatoriamente o nó AV, mas devido ao seu período refratário, nem todo impulso elétrico é propagado para os ventrículos. Com uma propagação caótica dos impulsos elétricos, os potenciais elétricos de diferentes regiões do átrio por vezes se cancela, o que resulta na ausência de ondas P, mas pode apresentar ondas fibrilatórias (ondas “f”), que são referentes às contrações descoordenadas do átrio. As ondas “f” apresentam menor amplitude do que as ondas P e nem sempre são seguidas de complexos QRS (SÖRNMO et al., 2018). A atividade elétrica caótica no átrio também causa descompasso no ritmo do ventrículo, pois, após seu período refratário, o nó AV dispara quando um próximo impulso elétrico de intensidade suficiente o atingir, o que ocorre de forma aleatória.

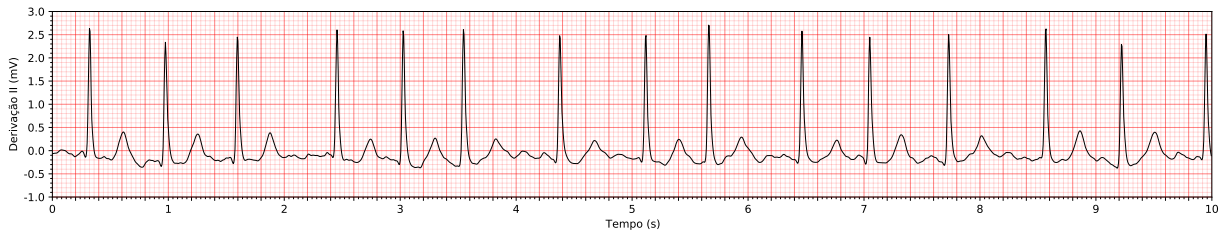
A Figura 16 mostra um exemplo do ritmo de um paciente com FA. Neste exemplo notamos a ausência de ondas P, sendo que alguns ciclos (após o 7º, 9º e 12º complexo

Figura 15: Vias de propagação dos impulsos (A) na fibrilação atrial e (B) no flutter atrial.



Fonte: Adaptado de (GUYTON; HALL, 2011)

Figura 16: Derivação II do ECG de um paciente com ritmo FA.



Fonte: Amostra obtida do banco de dados publicado por Reyna et al. (2021) (ID A0017).

QRS) apresentam ondas “f”. Outro fato diagnóstico é a variabilidade no intervalo R-R, indicando a assincronia entre átrios e ventrículos.

A FA aumenta o risco do desenvolvimento de diversas complicações. Estas complicações incluem condições leves, como fadiga, alterações na pressão arterial, tonturas, desmaios, cansaço, palpitações e dor no peito, e condições mais graves, como insuficiência cardíaca e formação de trombos (aumentando a possibilidade de ataque isquêmico transitório, embolia pulmonar e AVE) (REIFFEL, 2022; MARROUCHE et al., 2018; ZIKI et al., 2017). O tratamento da FA envolve principalmente o gerenciamento do risco de desenvolvimento de trombos (utilizando anticoagulantes), controle do ritmo cardíaco (utilizando drogas antiarrítmicas ou cardioversão elétrica) e redução de comorbidades que contribuem com a FA (obesidade, hipertensão e apneia do sono) (REIFFEL, 2022; STEFIL; LIP, 2022).

A detecção precoce da FA possibilita o início também precoce de terapias de controle de ritmo o que, segundo Kirchhof et al. (2020), reduz o risco do desenvolvimento de condições cardiovasculares adversas. Como em alguns casos esta arritmia pode se apresentar de forma assintomática, o telemonitoramento contínuo de pacientes de alto grau de risco

de desenvolvimento de FA (como aqueles no pós-operatório de cirurgias cardíacas) pode ser um elemento essencial para o diagnóstico e tratamento precoce.

### 2.1.2.5 Flutter atrial

O Flutter Atrial (FLU) é uma arritmia que, de forma similar à FA, gera um descompasso entre contrações atriais e ventriculares (GUYTON; HALL, 2011). Nesta arritmia, o sinal elétrico se propaga de forma circular ao longo dos átrios, devido a caminhos de baixa velocidade de condução que causam macrorreentradas sustentadas deste sinal ao longo do anel da válvula tricúspide (COSÍO et al., 2006; SAWHNEY; FELD, 2008; SAWHNEY et al., 2009; TAI; CHEN, 2009). Esta é uma arritmia que pode causar condições graves, tais como formações de trombos, isquemias e infartos (SAWHNEY; FELD, 2008).

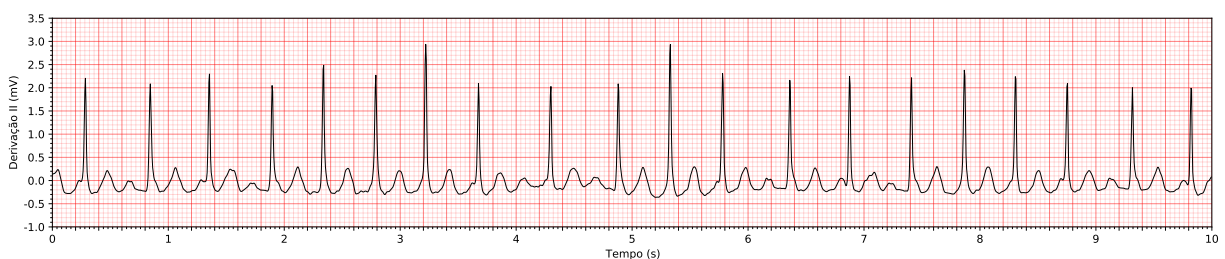
A Figura 15(B) exemplifica como a propagação do impulso elétrico se diferencia do impulso durante uma FA. Esta forma de propagação leva a uma taquicardia atrial (podendo atingir mais de 200 bpm), sendo que enquanto uma parte do átrio está contraída, a porção oposta está relaxada, e assim um baixo volume de sangue é bombeado pelo átrio (GUYTON; HALL, 2011).

A Figura 17 mostra um exemplo do ritmo de um paciente com FLU. Uma característica marcante do ritmo FLU no ECG é a presença de ondas “F”, que são padrões em dente de serra gerados pela circulação do impulso elétrico no átrio (COSÍO et al., 2006).

Devido ao período refratário do nó AV, a taquicardia atrial não é propagada para os ventrículos, sendo que ocorre um bloqueio de razão 2:1 ou de forma variável na condução dos estímulos para os ventrículos (SAWHNEY; FELD, 2008).

Na Figura 17 é possível notar que existe uma ausência de ondas P e as ondas “F” são de baixa amplitude. Desse modo, um sinal com maior nível de ruído pode ser confundido com um ritmo FA. Do ponto de vista de monitoramento contínuo de ritmo, esta possível

Figura 17: Derivação II do ECG de um paciente com ritmo FLU.



Fonte: Amostra obtida do banco de dados publicado por Reyna et al. (2021) (ID E00325).

confusão não é tão crítica, visto que as estratégias de tratamento de ambas arritmias é similar, e em alguns casos um paciente com FA pode ter seu ritmo convertido para FLU devido ao uso de drogas antiarrítmicas (CHEN, 2012).

Devido à similaridade do sinal de ECG (considerando a possibilidade de sinais ruidosos no contexto de dispositivos vestíveis) da FA e FLU, neste trabalho ambas arritmias serão agrupadas em uma única classe. Nas demais seções desta dissertação, a menos que especificado o contrário, o termo FA estará se referindo a ambas as arritmias.

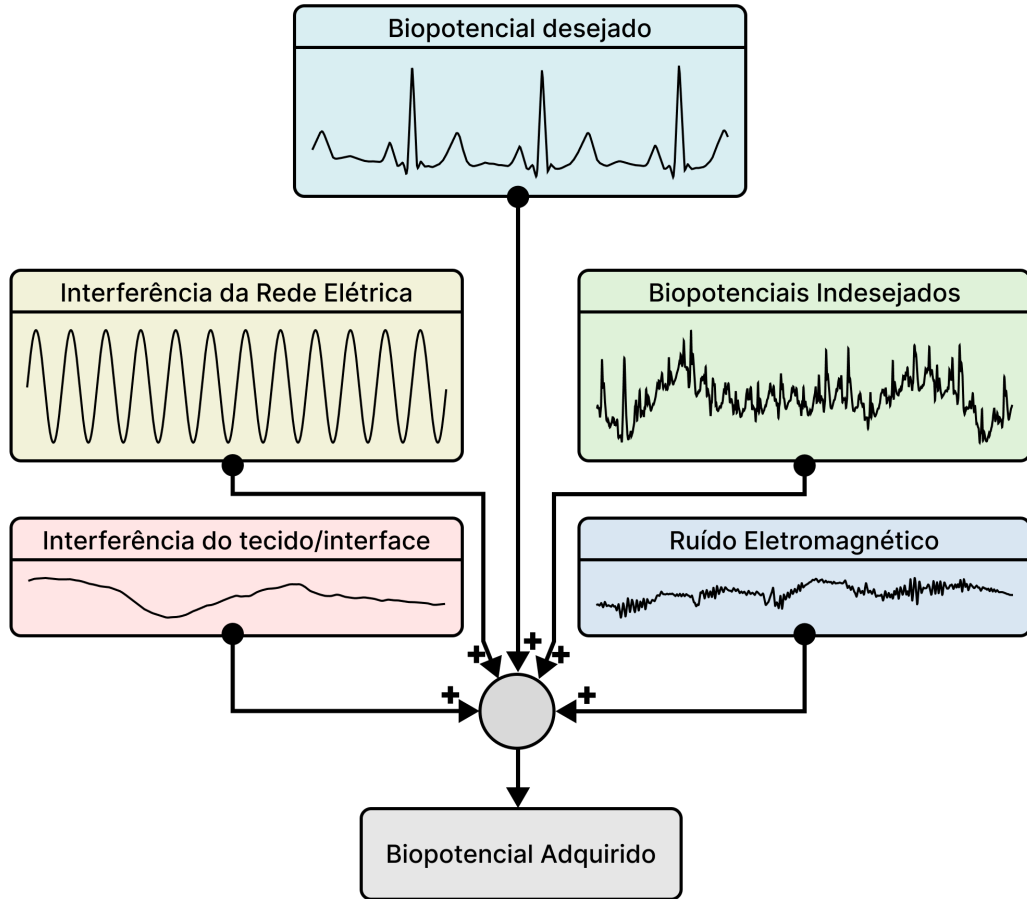
### 2.1.3 Sistemas de aquisição

A aquisição de sinais de ECG é realizada por meio de um eletrocardiógrafo. Este é um dispositivo médico capaz de amplificar, condicionar e registrar ao longo do tempo o sinal de diferença de potencial entre dois ou mais eletrodos posicionados no corpo de uma pessoa.

Por se tratar essencialmente de um voltímetro para tensões na ordem de micro a milivolts, é importante que as fontes de interferência externa sejam atenuadas. A Figura 18 mostra algumas fontes de interferência que compõem a diferença de potencial presente entre os eletrodos de um ECG.

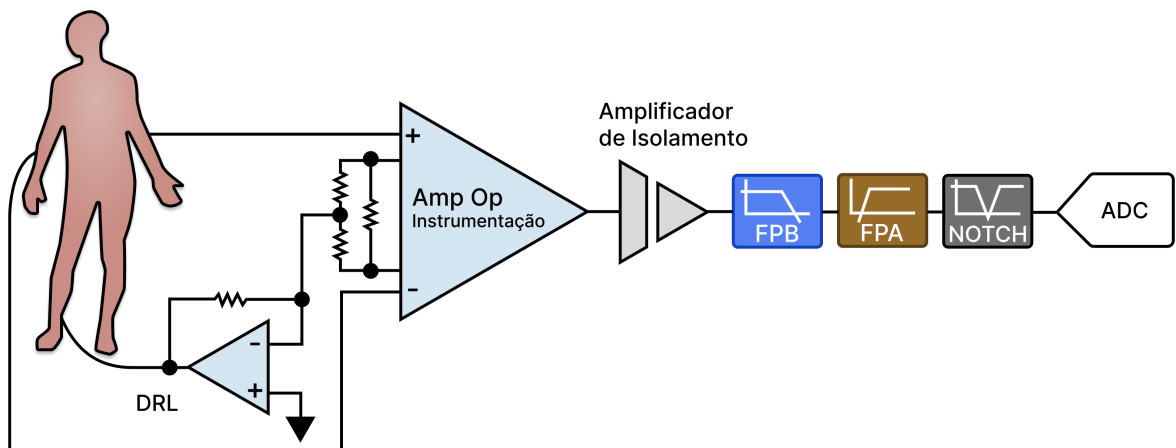
O condicionamento do sinal requer uma amplificação sem distorção do sinal, fornecendo uma alta impedância de entrada para reduzir a interferência do próprio eletrocardiógrafo na tensão a ser mensurada e utilizando amplificadores operacionais de precisão. A redução de interferências externas é realizada principalmente aumentando a rejeição de modo comum (utilizando topologias de circuito específicas para esta finalidade) e filtros (analógicos e digitais). A rejeição de modo comum atua em sinais que estão presentes nos dois eletrodos envolvidos na medida de diferença de potencial elétrico, como interferências eletromagnéticas e o nível DC inerente da interface eletrodo-pele. Alguns dispositivos ainda contam com rejeição de modo comum ativa, por meio de um circuito conhecido como *Driven Right Leg*, que injeta um sinal invertido correspondente à média entre os dois eletrodos de entrada em um eletrodo conectado à perna direita (por convenção, quando presente, este eletrodo é o terra do dispositivo) (SALINET; SILVA, 2019). Filtros passa-baixas, passa-altas e rejeita-faixa são utilizados, respectivamente, para reduzir interferências eletromagnéticas e biopotenciais de alta-frequência, para diminuir desvios de linha de base devido a movimentações do corpo, e eliminar interferência eletromagnética da rede. Após o condicionamento do sinal, o mesmo é digitalizado. A Figura 19 mostra o esquema de um típico eletrocardiógrafo de uma derivação.

Figura 18: Componentes do sinal de ECG medido.



Fonte: Adaptado de Salinet e Silva (2019), com sinais de Moody et al. (1984).

Figura 19: Circuito típico de um eletrocardiógrafo de uma derivação.

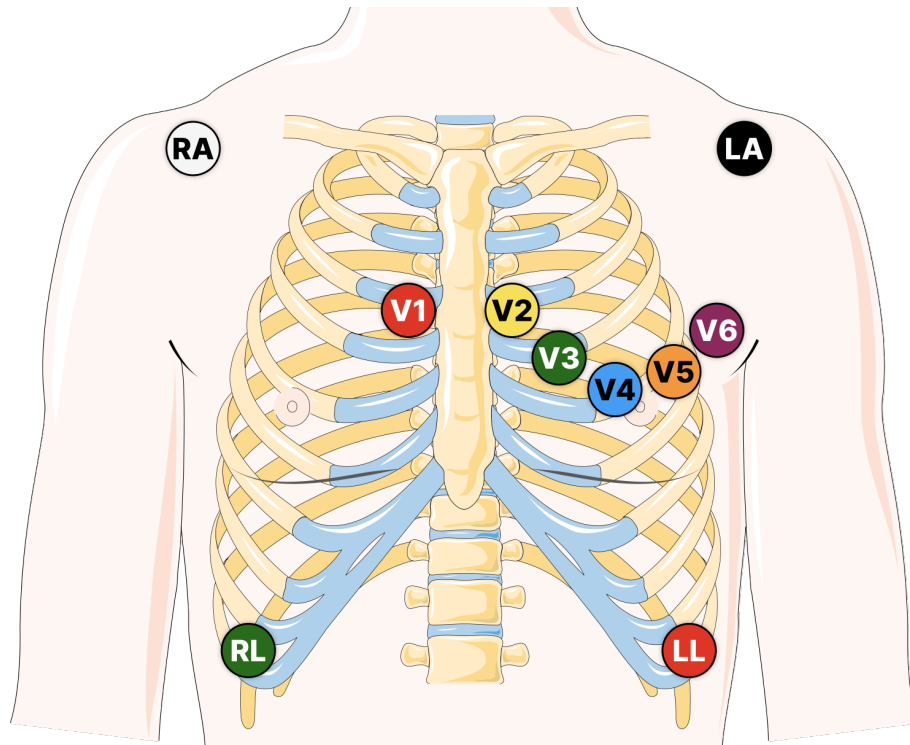


Fonte: Adaptado de Company-Bosch e Hartmann (2003).

O posicionamento dos eletrodos segue diferentes padrões, com diferentes números de eletrodos e sinais. O padrão mais comum é o de 12 derivações, o qual é exibido na Figura 20, e necessita de 10 eletrodos. As derivações são agrupadas em três tipos: de membros, aumentadas e precordiais. As derivações de membros são I, II e III, e as derivações aumentadas são aVL, aVR e aVF. Estas seis derivações são obtidas a partir de apenas três eletrodos e quatro destas derivações são combinações lineares das duas restantes, conforme equações a seguir (MOEINZADEH et al., 2020):

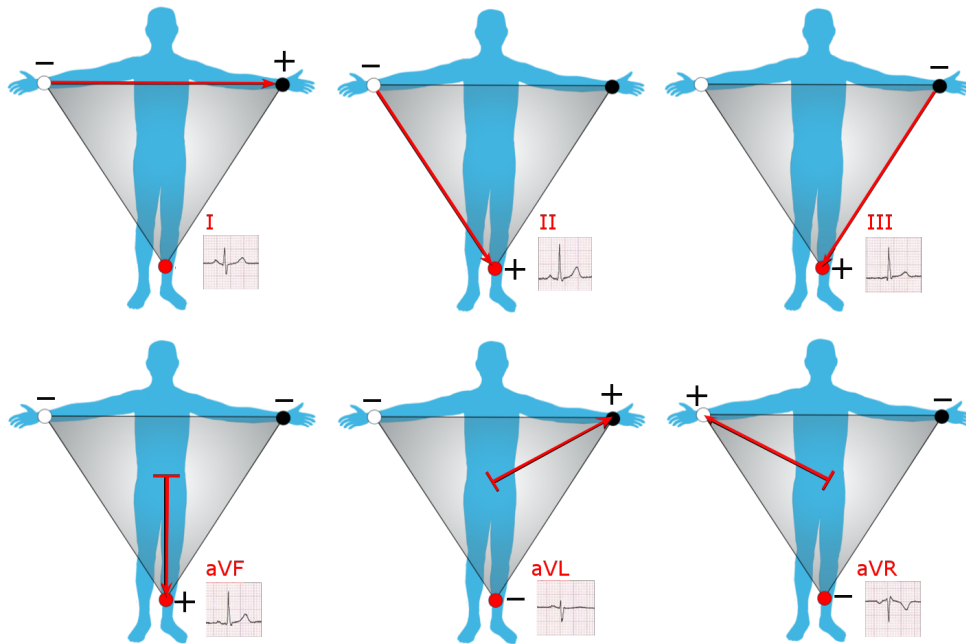
$$\begin{aligned}
 I &= LA - RA \\
 II &= LL - RA \\
 III &= LL - LA = II - I \\
 aVL &= LA - \frac{RA + LL}{2} = I - \frac{II}{2} \\
 aVR &= RA - \frac{LA + LL}{2} = -\frac{I + II}{2} \\
 aVF &= LL - \frac{RA + LA}{2} = II - \frac{I}{2}
 \end{aligned}$$

Figura 20: Posicionamento dos eletrodos na configuração de ECG de 12 derivações.



Fonte: Adaptado de SERVIER MEDICAL ART (2016) e Cadogan (2022).

Figura 21: Triângulo de Einthoven obtido com os eletrodos dos membros.



Fonte: Npatchett (2021).

Estas seis derivações podem ser analisadas por meio do triângulo de Einthoven (Figura 21), em que cada derivação é associada a diferentes direções no plano frontal do paciente. Como a aquisição destas seis derivações pode ser realizada com a amostragem de apenas dois sinais, ao longo deste texto os termos “ECG de seis derivações” e “ECG de dois sinais” serão utilizados como sinônimos, visto que o processamento do ECG ocorrerá apenas nos dois sinais independentes, mas do ponto de vista clínico o exame é de seis derivações.

As derivações precordiais (V1 a V6) são mensuradas como a diferença de potencial entre os eletrodos correspondentes e uma referência chamada terminal central de Wilson, o qual é definido como a média entre os potenciais elétricos dos eletrodos LA, RA e LL (GARGIULO, 2015).

#### 2.1.4 Dispositivos vestíveis

Na forma de monitor Holter, os eletrocardiógrafos vestíveis já são utilizados na cardiologia desde 1962 (MAR, 2005). Estes dispositivos permitem a aquisição contínua dos sinais durante longos períodos de tempo. Um exame de ECG de 12 derivações tradicional registra 10 segundos da derivação II e cerca de 2,5 segundos das demais derivações, já um exame de Holter considerado de curta duração pode registrar entre 24 a 48 horas de sinal de todas as derivações disponíveis e um Holter de longa duração pode armazenar de uma



a duas semanas de sinais (ZIMETBAUM; GOLDMAN, 2010; GALLI et al., 2016).

Novas tecnologias de eletrocardiógrafos vestíveis também estão disponíveis na forma de monitores de eventos (*loopers* e pós-evento) e telemetria móvel.

Os monitores de evento são dispositivos semelhantes ao Holter tradicional, com o diferencial de armazenar apenas os trechos de sinal referentes a eventos cardíacos (ZIMETBAUM; GOLDMAN, 2010). Os monitores de evento do tipo *looper* possuem uma memória específica que armazena uma janela móvel do sinal, sendo que esta janela pode ser armazenada na memória principal de forma automática (o sinal da janela é continuamente processada por algoritmos de detecção de eventos cardíacos) ou manual (o paciente dispara a gravação ao sentir algum sintoma); já os monitores de pós-evento não possuem uma gravação contínua dos sinais e armazena apenas trechos de sinal após o paciente disparar a gravação (GALLI et al., 2016). Monitores de evento externos são usualmente utilizados em exames com durações de até um mês, já os monitores implantáveis podem ser utilizados por até três anos (GALLI et al., 2016).

A telemetria móvel é uma variante dos monitores de eventos que possuem a capacidade de monitoramento remoto, por meio do envio dos sinais de ECG para servidores na nuvem. Nesta modalidade o eletrocardiógrafo usualmente envia os sinais por Bluetooth para um smartphone, o qual pode retransmitir o sinal como um todo em tempo real ou pode atuar como um *looper*, enviando os sinais apenas quando um possível evento for detectado.

Também é possível dividir a eletrocardiografia móvel nas categorias de diagnóstico manual ou automático. A análise de exames Holter e monitores de evento é primordialmente realizada de forma manual, podendo ser assistida por algoritmos. Neste caso os algoritmos são utilizados para destacar trechos com potenciais eventos e o médico determina qual é o diagnóstico, sendo que em exames de Holter o médico pode inclusive analisar manualmente o sinal como um todo sem a assistência de algoritmos. A telemetria móvel é uma tecnologia mais recente e que geralmente incorpora algoritmos não apenas de detecção de eventos mas também de classificação, gerando um diagnóstico automático. O diagnóstico gerado pelos algoritmos é posteriormente revisado por um médico que é responsável pelo diagnóstico final. A diferença entre estas duas categorias está no grau de automaticidade do diagnóstico, visto que mesmo o diagnóstico automático ainda deve ser revisado manualmente por um médico (WILLCOX et al., 2021).

#### 2.1.4.1 Desafios técnicos

O uso de eletrocardiógrafos vestíveis apresenta alguns desafios técnicos adicionais em relação aos dispositivos tradicionais. Como estes eletrocardiógrafos são alimentados por bateria, o consumo energético deve ser reduzido ao mínimo possível, para prolongar o período de monitoramento contínuo. Outro desafio é a presença de um maior nível de ruídos, devido a características dos eletrodos próprios para estas aplicações e à movimentação inerente do registro de biopotenciais em dispositivos vestíveis (RAHMAN et al., 2022).

A redução no consumo energético é geralmente obtida por otimizações de hardware, onde existe uma relação de compromisso entre largura de banda do sinal e frequência de amostragem: maior frequência de amostragem permite a aquisição de detalhes de alta frequência, porém exige um maior consumo (BOUZID et al., 2022). Algumas aplicações, em especial na telemetria móvel, também podem ser otimizadas quanto ao protocolo de comunicação e processamentos embarcados. Um caso extremo de otimização energética é o monitor pós-evento, pois o dispositivo permanece desligado até o momento que o paciente inicie a coleta. A telemetria móvel intermediada por smartphone também fornece uma otimização energética, visto que o eletrocardiógrafo realiza a aquisição e transmite os sinais por protocolos energeticamente eficientes, como *Bluetooth Low Energy*, a transmissão para a nuvem (e possivelmente alguns processamentos) é realizada no smartphone e processamentos mais complexos são executados nos servidores em nuvem. Ainda no contexto da telemetria móvel, a eficiência energética dos aplicativos de smartphone responsáveis pelo processamento local e retransmissão dos sinais também é desejável, mesmo que não obrigatória. Desse modo, algoritmos leves e robustos para a classificação do ritmo cardíaco podem ser um diferencial em um telemonitoramento de arritmias.

Ruídos sempre estão presentes em qualquer tipo de aquisição de sinais. Sinais de biopotenciais, como o ECG, ainda sofrem da presença de diversos outros biopotenciais contaminantes, como eletromiogramas e flutuações de potencial devido a alterações na impedância do corpo. Aplicações com eletrocardiógrafos vestíveis devem ser capazes de lidar com níveis muito elevados de ruídos e artefatos, em especial aqueles de origem no movimento, visto que a aquisição é realizada enquanto o paciente está em suas atividades cotidianas. Alguns eletrocardiógrafos vestíveis ainda apresentam um desafio adicional no quesito de ruídos, pois alguns dispositivos utilizam eletrodos secos (como é o caso da maioria dos *smartwatches* com capacidade de ECG) e sofrem de alta impedância de contato, aumentando o ruído ao mesmo tempo que atenuando o sinal, além de apresentar

maior acoplamento de modo comum e nível DC (UDUPA, 2021).

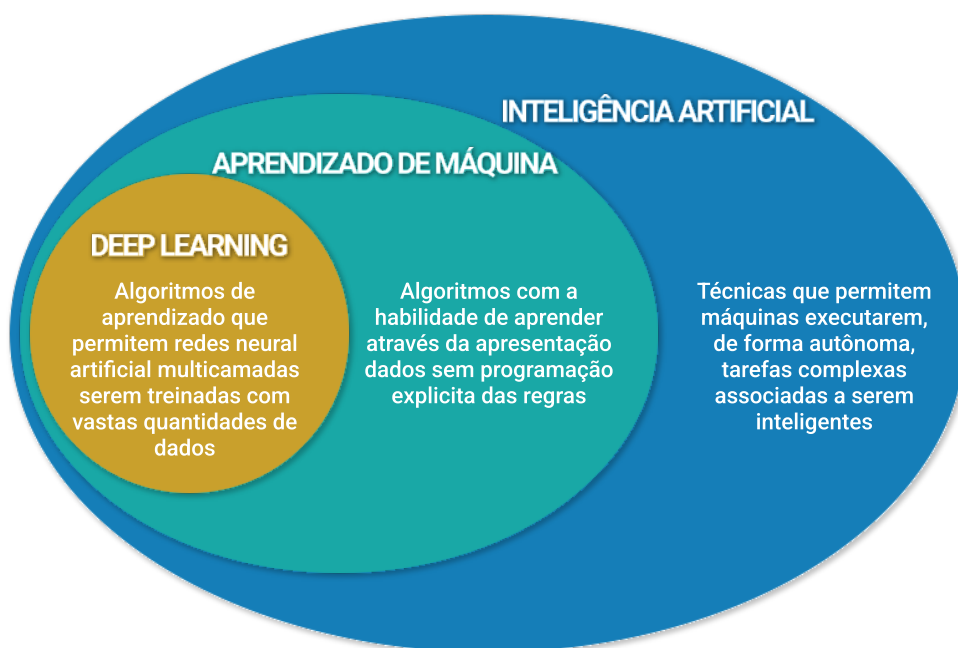
Eletrocardiógrafos vestíveis do tipo “*patch*” geralmente possuem todos os eletrodos em um único adesivo e a distância entre eletrodos é pequena. Nestes casos é possível diminuir os ruídos e artefatos de movimentação ao posicionar o *patch* em posições biomecanicamente mais estáveis, como o mediastino. Por outro lado, eletrocardiógrafos que utilizam fios para conectar os eletrodos aos membros são suscetíveis não apenas à movimentação do pacientes, mas também à movimentação dos próprios fios e a efeitos de acoplamento eletromagnético com fontes ambientais externas.

O processamento do ECG em aplicações vestíveis sempre deve considerar a qualidade do sinal adquirido, sendo por meio de pré-processamentos específicos para atenuação de ruídos ou por descarte de trechos com baixa qualidade do sinal (BOUZID et al., 2022).

## 2.2 Aprendizado de máquina

A inteligência artificial é um campo da Ciência da Computação que estuda o desenvolvimento de algoritmos e técnicas que permitem um computador realizar tarefas complexas, normalmente atribuídas a seres inteligentes (Figura 22). Estas técnicas englobam desde algoritmos simples, como sistemas baseados em regras, até sistemas complexos, capazes

Figura 22: Diagrama de Venn associando inteligência artificial, aprendizado de máquina e *deep learning*.



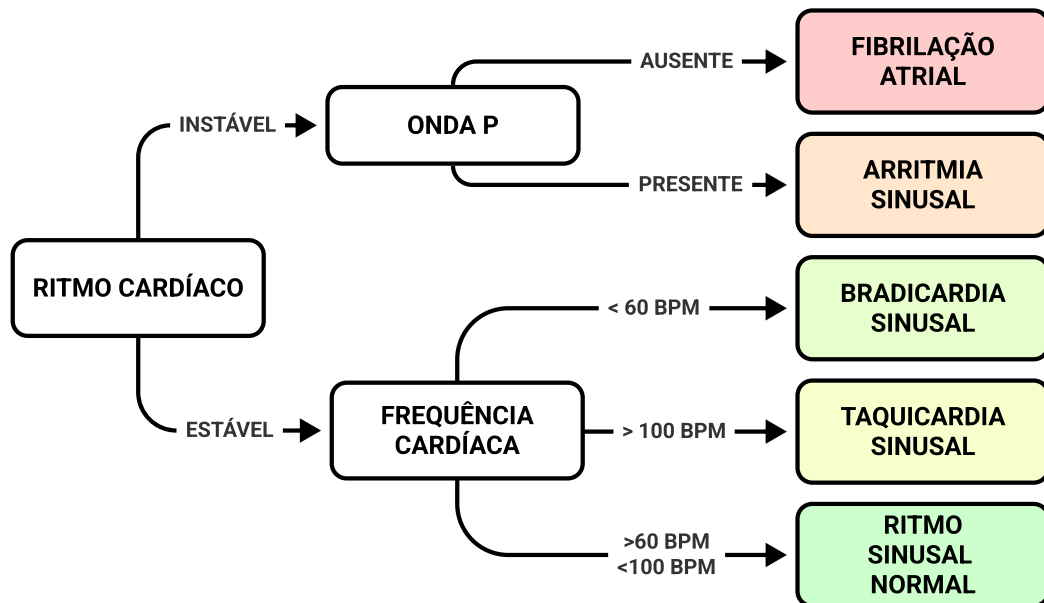
Fonte: De autoria própria.

de aprender automaticamente a partir de interações com o ambiente.

Um exemplo de algoritmo de inteligência artificial é a árvore de decisões. A Figura 23 apresenta um exemplo de árvore de decisão para classificação de ritmo cardíaco. Neste exemplo o algoritmo possui três entradas (estabilidade do ritmo cardíaca, presença de ondas P e frequência cardíaca) e uma série de comparações lógicas geram a saída. As regras presentes nesta árvore podem ser determinadas manualmente por um especialista ou de forma automática por algoritmos de aprendizado de máquina, a partir de exemplos de entradas e saídas. Em ambientes bem controlados e de simples determinação dos possíveis cenários é possível utilizar regras criadas manualmente, este é o caso da inteligência artificial presente em diversos jogos eletrônicos. Já situações complexas, com cenários por vezes imprevisíveis, como identificação de spam, reconhecimento de objetos em imagens e direção de carros autônomos, torna-se inviável determinar todas as regras de forma manual. Neste contexto é necessário utilizar de algoritmos de aprendizado de máquina associados com grandes quantidades de dados de treinamento.

O aprendizado de máquina é um subcampo da inteligência artificial focado no aprendizado automático. Para este aprendizado são construídos modelos computacionais com métricas de avaliação associadas. O objetivo dos algoritmos de aprendizado de máquina é otimizar as métricas dos modelos com base em um conjunto de dados de treinamento.

Figura 23: Exemplo de árvore de decisão.



Fonte: De autoria própria.

Alguns exemplos de aprendizado de máquina incluem:

- Regressão linear: Utilizando a equação  $f(X) = AX + \beta$  como modelo e um vasto conjunto de amostras, técnicas com mínimos quadrados podem ser utilizadas para aprender o hiperplano que melhor otimiza o erro quadrático entre inferência do modelo e dados de treino;
- Máquina de vetores de suporte: Utiliza funções que mapeiam a entrada para um novo hiperespaço onde são traçados hiperplanos, de forma a maximizar a margem entre duas ou mais classes;
- Árvores de decisão: Utiliza como modelo uma árvore binária constituída de comparações sucessivas. Esta árvore pode ser montada de forma automática otimizando o ganho de informação (diminuindo a entropia) em cada nó da árvore;
- Redes neurais artificiais: Utilizam modelos originalmente inspirados nos neurônios biológicos. Matematicamente, estes modelos são constituídos de multiplicações matriciais e ativações não lineares. As métricas a serem otimizadas são problema dependentes e a otimização é realizada por meio de métodos de retropropagação de erros.

As técnicas de aprendizado de máquina podem ser categorizadas basicamente em supervisionadas e não supervisionadas. As técnicas supervisionadas necessitam de exemplos rotulados, ou seja, pares de entradas e saídas, enquanto as técnicas não supervisionadas são capazes de aprender a partir de dados não rotulados. O aprendizado supervisionado geralmente é utilizado nas tarefas de classificação e regressão, de acordo com o tipo de saída do modelo, respectivamente saída discreta (categórica) e contínua. Uma das principais aplicações do aprendizado não supervisionado é a clusterização, ou seja, a agregação de objetos em grupos com característica similares.

### 2.2.1 Redes Neurais Artificiais

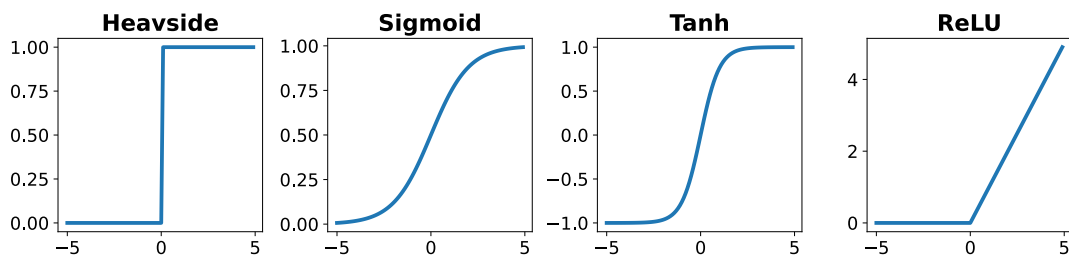
As redes neurais artificiais, ou simplesmente redes neurais, são modelos de aprendizado de máquina que possuem inspiração original no neurônio biológico. As redes neurais são constituídas de unidades, ou neurônios, conectados entre si em forma de camadas. A nomenclatura neurônio artificial advém da proposta original feita por McCulloch e Pitts (1943), onde os neurônios são modelados com uma característica de “tudo ou nada”. Nesta modelagem o neurônio recebe diversas entradas e o neurônio é ativado apenas se a

soma ponderada por determinados pesos é superior a um limiar. As redes neurais atuais utilizam unidades que são generalizações deste neurônio de McCulloch e Pitts, sendo matematicamente modeladas como:

$$y = \varphi \left( w_0 + \sum_{i=1}^m w_i x_i \right) \quad (2.1)$$

em que  $x_i$  são as entradas da unidade,  $w_i$  são os pesos,  $w_0$  é o limiar ou bias da unidade e  $\varphi$  é uma função de ativação. Utilizando a função de Heaviside como ativação, obtemos o neurônio artificial original. As principais funções de ativação utilizadas são: identidade, logística (ou sigmoid), tangente hiperbólica (tanh) e unidade linear retificada (ReLU). A Figura 24 apresenta a curva destas funções de ativação.

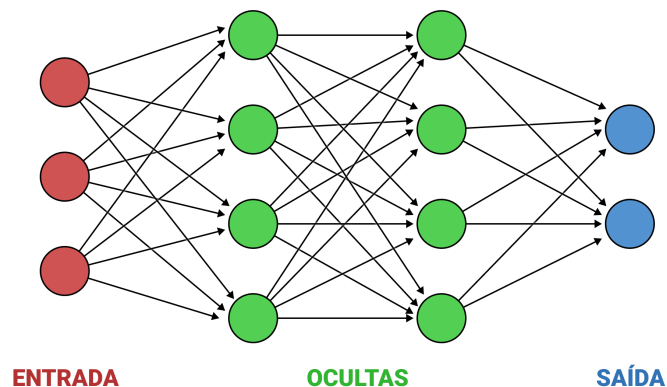
Figura 24: Exemplos de funções de ativação.



Fonte: De autoria própria.

As redes neurais são formadas por camadas de unidades neurais, sendo que a primeira camada recebe as entradas do modelo, as camadas ocultas utilizam a saída da camada anterior como entrada e a última camada retorna a saída do modelo. A Figura 25 exhibe o diagrama de uma rede neural artificial do tipo perceptron multicamada.

Figura 25: Diagrama de uma rede neural artificial do tipo perceptron multicamada.



Fonte: De autoria própria.

## 2.2.2 Deep learning

Os neurônios de McCulloch e Pitts não possuem habilidade de aprendizado a partir de dados, sendo que os pesos são configurados manualmente. Rosenblatt (1958) propôs uma adaptação ao neurônio de McCulloch e Pitts, chamada de perceptron, e uma regra de atualização dos pesos que permite o aprendizado automático a partir de amostras de entrada e saída. Utilizando o algoritmo de retropropagação aplicado a redes neurais proposto por Werbos (1974), Rumelhart et al. (1986) propôs as redes neurais do tipo perceptron multicamada.

O perceptron multicamada com pelo menos duas camadas ocultas já consiste de uma técnica de *deep learning*, visto que é um modelo que consiste de mais de uma camada de transformação não linear em sequência. No entanto, o termo *deep learning* por vezes é associado a redes recorrentes e redes multicamadas do tipo convolucional. Neste trabalho abordaremos apenas as redes profundas convolucionais.

## 2.2.3 Redes Neurais Convolucionais

Nas redes neurais convolucionais as camadas são compostas de filtros convolucionais, podendo ser procedidos de uma função de ativação. Portanto, a saída de uma camada convolucional, chamada mapa de ativação, consiste em aplicar uma operação de convolução entre um tensor de entrada e os filtros (ou pesos) da camada. Sendo mais matematicamente rigoroso, a operação aplicada na realidade é uma correlação cruzada, definida pela equação 2.2.

$$(X * K)[n] \triangleq \sum_{m=-\infty}^{\infty} K[m]X[m+n] \quad (2.2)$$

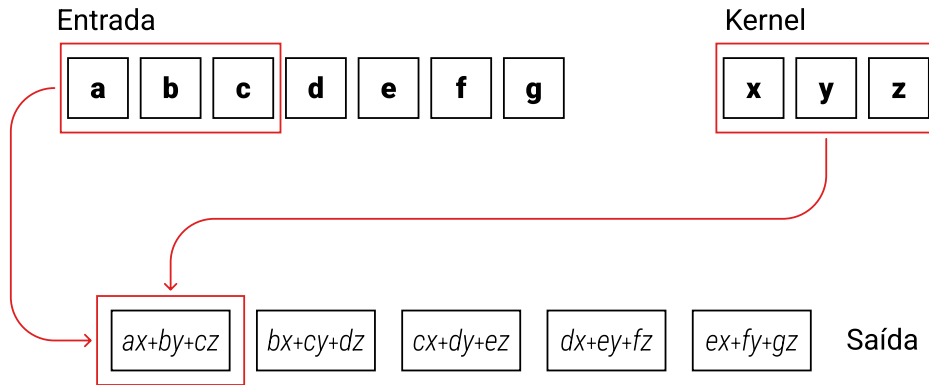
Ou generalizando para tensores com dimensionalidade M:

$$\left(X * K\right)[n_1, \dots, n_M] = \sum_{i_1=-\infty}^{\infty} \dots \sum_{i_M=-\infty}^{\infty} K[i_1, \dots, i_M]X[n_1 + i_1, \dots, n_M + i_M] \quad (2.3)$$

A Figura 26 exibe uma representação visual da convolução, que pode ser interpretada como um conjunto de pesos se deslocando ao longo do sinal de entrada e para cada passo do deslocamento uma posição da saída é determinada pela soma da entrada ponderada pelo filtro. Neste exemplo, foi considerada uma convolução sem padding, ou seja, considerando

apenas posições onde o kernel inteiro está dentro da entrada. Observa-se que sem padding, a convolução resulta em uma saída com menor dimensão. Para manter a mesma dimensão, considerando um kernel de tamanho  $2K + 1$ , é necessário adicionar  $K$  zeros no início e fim do vetor de entrada.

Figura 26: Visualização da operação de convolução 1D sem padding.



Fonte: De autoria própria.

Na Figura 26, foi representado o exemplo mais simples de convolução, onde temos apenas um filtro e a entrada possui apenas um canal. Entretanto, as entradas de uma camada convolucional geralmente são multicanais e estas camadas possuem vários filtros. Para este caso genérico, a saída da camada convolucional é dada pela equação 2.4.

$$(X * K)[n, c_o] \triangleq \sum_{c_i} \sum_m K[m, c_i, c_o] X[n + m, c_i] \quad (2.4)$$

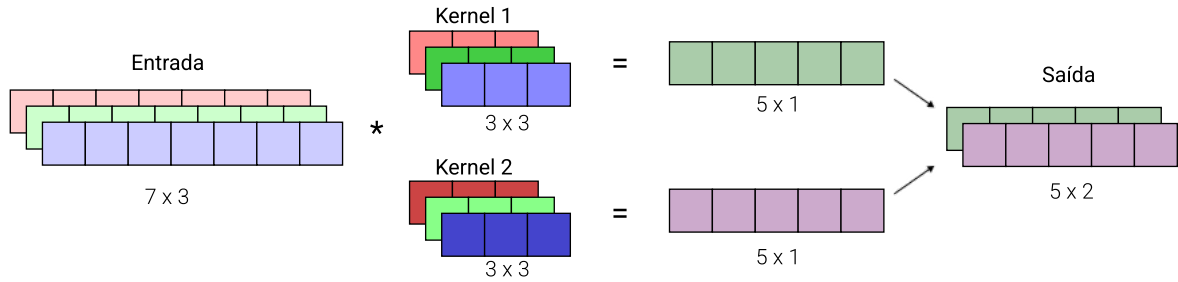
Onde  $X$  é uma matriz unidimensional de tamanho  $N$  e  $C_i$  canais, e  $K$  é o kernel da convolução com  $C_o$  filtros de tamanho  $M$ . A Figura 27 exibe uma visualização gráfica da convolução de um vetor com três canais, de tamanho 7, por um kernel com dois filtros de tamanho 3. Note que cada filtro do kernel possui o mesmo número de canais que a entrada.

Considerando uma camada convolucional com  $C_o$  filtros de tamanho  $M$ , que tem por entrada um mapa de ativação de tamanho  $N$ , com  $C_i$  canais, seu kernel terá um tamanho total de  $M \cdot C_i \cdot C_o$  parâmetros e necessitará de um total de  $M \cdot C_i \cdot C_o \cdot N$  operações de ponto flutuante (do tipo *multiply-accumulate*) para computar a sua saída (HOWARD et al., 2017).

O custo computacional e o número de parâmetros de uma camada convolucional tradicional aumenta proporcionalmente ao tamanho da entrada (dimensão e número de



Figura 27: Visualização da operação de convolução 1D para múltiplos filtros e entrada com múltiplos canais.



Fonte: De autoria própria.

canais) e ao tamanho do kernel (dimensão e número de filtros). Deste modo, esta operação se torna muito custosa para entradas ou kernels muito grandes. Uma alternativa que reduz este custo computacional é o uso de convoluções *depthwise* separável (SepConv) (CHOLLET, 2017).

Na convolução separável a operação de convolução é fatorada em duas convoluções mais simples. A primeira convolução, chamada *depthwise*, consiste de utilizar kernels com apenas um canal e aplicá-los independentemente para cada canal da entrada, gerando um mapa de ativações intermediário ( $X_{int}$ ), segundo a equação 2.5.

$$X_{int}[n, c_i] \triangleq \sum_m K[m, c_i] X[n + m, c_i] \quad (2.5)$$

Este mapa de ativação intermediário então passa por uma convolução tradicional, mas com kernel de comprimento unitário, segundo a equação 2.6.

$$X_{out}[n, c_o] \triangleq \sum_{c_i} K[c_i, c_o] X[n, c_i] \quad (2.6)$$

Das equações 2.5 e 2.6, uma operação de convolução separável necessita de  $M \cdot C_i + C_i \cdot C_o$  parâmetros em seus kernels e tem um custo computacional de  $N \cdot (M \cdot C_i + C_i \cdot C_o)$  operações de ponto flutuante. Comparando com o custo computacional da convolução tradicional, observa-se uma redução por um fator de  $\frac{1}{C_o} + \frac{1}{M}$  no total de operações e de parâmetros da camada.

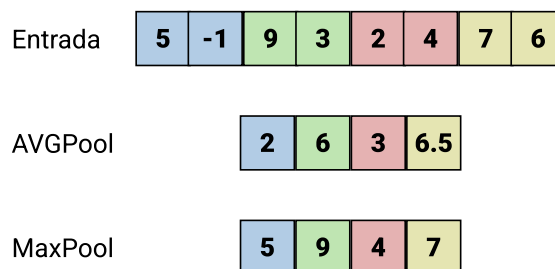
Uma alternativa à camada de convolução separável é uma camada puramente *depthwise* (DwConv). Como por padrão a operação de convolução *depthwise* gera uma saída com o mesmo número de canais que a entrada, esta camada conta com um fator de multiplicação

de profundidade  $dm$ , inteiro e maior que zero. Este fator define quantos filtros serão aplicados em cada canal de entrada. Uma camada com um fator  $dm > 1$  também pode ser interpretado como a concatenação de  $dm$  convoluções *depthwise* independentes, portanto tal camada possui  $M \cdot Ci \cdot dm$  parâmetros e necessita de  $N \cdot M \cdot Ci \cdot dm$  operações para realizar uma inferência.

Nas redes neurais convolucionais, comumente são utilizadas estratégias para reduzir o tamanho do mapa de ativação gerado por algumas camadas da rede, com o intuito de reduzir o custo computacional das camadas a seguir. Dentre estas estratégias, temos o uso de *stride* na convolução e o uso de camadas de *pooling*. O *stride* indica o passo da convolução. Aplicando este conceito na equação 2.4 e considerando passo 3, seria o equivalente a aplicar a equação apenas para  $n \in [0, 3, 6, 9, \dots, 3i]$ . As operações de *pooling* também envolvem o conceito de *stride*, sendo que uma janela deslizante de determinado tamanho percorre o sinal de entrada com um determinado *stride* e para cada passo desta janela, uma função de redução é aplicada. As operações de redução mais comuns são média e valor máximo, dando origem às operações de *average pooling* (AVGPool) e *max pooling* (MaxPool), respectivamente. A figura 28 mostra um exemplo da aplicação destas operações de *pooling*, considerando *stride* e janela móvel de tamanho 2, em um vetor 1D. Tanto o uso de *stride* na convolução como o uso de operações de *pooling* produzem uma saída que terá suas dimensões reduzidas por um fator  $1/stride$ . As operações de *pooling* não alteram o número de canais do mapa de ativação, pois são aplicadas apenas ao longo do eixo do comprimento do sinal.

A princípio, uma rede neural completamente convolucional pode ser utilizada para diversas tarefas, em especial tarefas que envolvem um sinal de entrada e outro de saída. Entretanto, na tarefa de classificação de um sinal 1D, a saída consiste de um ou mais

Figura 28: Exemplo de operações de AVGPool e MaxPool, considerando *stride* e janela de tamanho 2.



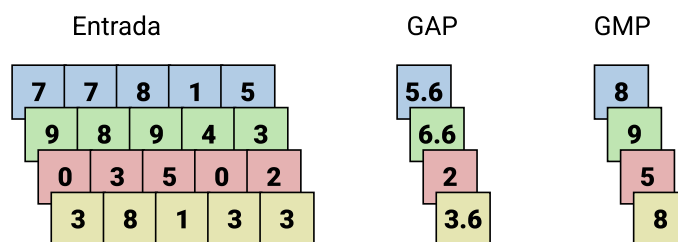
Fonte: De autoria própria.

rótulos. Neste caso, usualmente o modelo de classificação consiste de um bloco convolucional seguido por um bloco de perceptron multicamadas, que neste contexto também é chamado de bloco de camadas densas. Os blocos convolucionais unidimensionais possuem como entrada e saída uma matriz bidimensional (comprimento da entrada x número de canais), mas uma camada densa possui como entrada um vetor unidimensional. Para conectar um bloco convolucional a um bloco denso é necessário aplicar uma operação de achatamento.

Para as análises a seguir, será considerado o achatamento de um mapa de ativações  $X_{N \times C}$ , com tamanho  $N$  e  $C$  canais, que constituirá de entrada para uma camada densa com  $H$  unidades. A operação de achatamento mais simples consiste de concatenar os canais de uma matriz de forma a criar um vetor, de tamanho  $N \cdot C$ , contendo todos os valores da matriz. Por si própria, esta operação não possui alto custo computacional, entretanto seu uso pode acarretar em elevado número de parâmetros na próxima camada densa, pois a mesma possuirá  $N \cdot C \cdot H$  parâmetros. Por vezes, quando o mapa de ativação  $X_{N \times C}$  possui grandes tamanhos, uma redução de informação pode ser aplicada para reduzir o número de parâmetros das camadas a seguir. Uma técnica amplamente utilizada é o *pooling* global (LIN et al., 2014), onde uma função de redução (geralmente valor máximo ou média) é aplicada ao longo de cada canal do mapa de ativação, resultando em um vetor de saída de tamanho  $C$ . A Figura 29 ilustra os processos de *pooling* globais quando aplicados a um mapa de ativação de tamanho  $5 \times 4$ . O uso de *pooling* global tem maior impacto quanto maior for o mapa de ativação, mas deve ser utilizado com precaução quando sua entrada possuir poucos canais, já que ocorrerá uma redução brusca na informação propagada pela rede.

As redes neurais convolucionais podem contar ainda com outras camadas diversas. Aqui destacamos as camadas de normalização por lotes e de *dropout*, sendo responsáveis,

Figura 29: Exemplo de operações de *Global Average Pooling* (GAP) e *Global Max Pooling* (GMP).



Fonte: De autoria própria.

respectivamente, por acelerar o aprendizado e por gerar efeito de regularização.

A normalização por lotes, em inglês *batch normalization*, consiste de aplicar uma normalização z-score no mapa de ativação de cada lote de treinamento (IOFFE; SZEGEDY, 2015). No entanto, esta normalização é realizada por meio de parâmetros de média e variância ajustados durante o treinamento. Para cada lote apresentado à rede durante o treinamento, são calculadas estas medidas estatísticas para o mapa de ativação e o valor calculado é utilizado para atualizar os parâmetros da camada por meio de uma regra de atualização (como, por exemplo, uma regra de *momentum*). Desse modo, durante a inferência, os mapas de ativação são normalizados de acordo com os parâmetros finais do treinamento. O intuito desta técnica é manter a variância dos mapas de ativação próximas ao valor unitário, evitando o problema de desvanecimento do gradiente. Com isso, a propagação do gradiente das perdas da rede é estabilizado e torna-se possível treinar redes mais profundas com um número menor de épocas.

Camadas de *dropout* são utilizadas como regularização, para diminuir os efeitos de *overfitting*, ou seja, evitar que o modelo esteja tão bem ajustado aos dados de treinamento que não seja capaz de generalizar para novos dados (SRIVASTAVA et al., 2014). O *dropout* tem efeito apenas durante o treinamento e consiste em zerar aleatoriamente, com uma probabilidade  $p$ , os elementos de um mapa de ativação. Isto faz com que a cada passo de treinamento sejam atualizados um conjunto diferente de pesos (já que ao zerar um mapa de ativação, o gradiente para de ser propagado), tornando menos provável que a rede se torne ajustada apenas aos dados de treinamento. Uma variante desta camada é o *dropout* espacial, onde canais inteiros do mapa de ativação são zerados (TOMPSON et al., ). É importante notar que a ação de zerar alguns elementos tem efeito de escalonar os mapas de ativação desta camada por um fator de  $(1 - p)$ . Para compensar este efeito, o mapa de ativação é escalonado por um fator  $\frac{1}{1-p}$ . Durante a fase de inferência nenhum escalonamento é necessário, pois nenhum elemento é zerado.

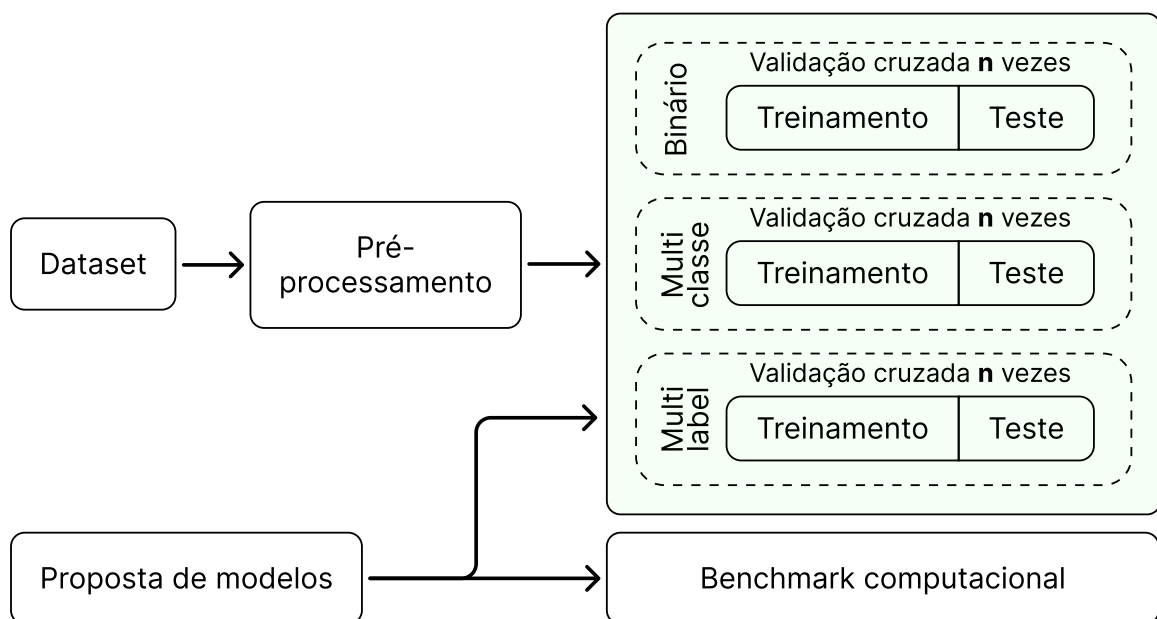
Nos últimos anos, a implementação de modelos de redes neurais convolucionais tem se tornado mais simples. Frameworks como o TensorFlow (ABADI et al., 2016), Keras (CHOLLET et al., 2015) e PyTorch (PASZKE et al., 2019), implementam algoritmos que realizam a retropropagação de erros necessária para o treinamento por gradientes descendentes, além de implementar diversas técnicas de otimização, perdas, regularizadores e blocos diversos. O acervo de ferramentas disponibilizadas por tais frameworks possibilitam que o trabalho de pesquisa seja mais focado em inovações, já que as bases para o design de modelos está disponível de uma forma bem consolidada.

### 3 METODOLOGIA

Neste capítulo será abordada a metodologia aplicada na pesquisa que deu origem a este texto. A Figura 30 sumariza, em alto nível, a metodologia utilizada na pesquisa. Modelos de aprendizado profundo foram propostos e foi realizado um benchmark para medir seus respectivos custos computacionais. Uma proposta de pré-processamento foi utilizada para condicionar os sinais de ECG anotados de um dataset público. Tais dados foram utilizados para avaliação da classificação dos modelos por meio de validação cruzada em experimentos de classificação binária, multiclasse e multilabel.

Neste capítulo se encontram: descrição do dataset utilizado, motivação e detalhamento dos pré-processamentos adotados, descrição dos modelos propostos (modelos autorais e adaptações de modelos clássicos), métricas utilizadas para análise da performance de classificação e detalhamento da implementação da classificação.

Figura 30: Esquemático de alto nível da metodologia do trabalho.



Fonte: De autoria própria.

### 3.1 Base de dados

Os sinais de ECG utilizados neste trabalho são provenientes da competição George B. Moody PhysioNet Challenge 2021 - Will Two Do? Varying Dimensions in Electrocardiography (REYNA et al., 2021). Esta competição é uma extensão da competição do ano 2020, que teve como tema e objetivo a identificação de diagnósticos clínicos a partir de sinais de ECG de 12 derivações. Em 2021, a competição aumentou o volume de sinais disponibilizados publicamente e manteve o mesmo objetivo de identificação de diagnósticos clínicos, entretanto não apenas com ECG de 12 derivações, mas também com conjuntos reduzidos de derivações, com o caso extremo sendo o diagnóstico com apenas as derivações I e II.

Este desafio compilou um conjunto de datasets com um total de mais de 130 mil exames, sendo que deste total 88 253 foram disponibilizados publicamente. A seguir são listados cada uma das fontes dos datasets individuais que compõem o conjunto de dados disponibilizado pelo desafio:

- **China Physiological Signal Challenge - CPSC2018** (LIU et al., 2018): Com um total de 13 256 exames, 10 330 foram disponibilizados publicamente. Os sinais foram amostrados em 500 Hz e possuem duração variada, entre 6 e 144 segundos.
- **St Petersburg INCART** (GOLDBERGER et al., 2000): 74 sinais de ECG retirados de 32 exames Holter. Cada sinal possui 30 minutos e foi adquirido a 257 Hz.
- **Physikalisch-Technische Bundesanstalt**: dois datasets desta fonte:
  - *PTB* (BOUSSELJOT et al., 1995): 516 exames com duração de até 120 segundos, amostrados a 1000 Hz.
  - *PTB-XL* (WAGNER et al., 2020): 21 837 exames com duração de 10 segundos e 500 Hz de frequência de amostragem.
- **Georgia**: Com um total de 20 672 exames, 10.344 foram disponibilizados publicamente. Exames com 10 segundos de duração, amostrados a 500 Hz.
- **Chapman University, Shaoxing People's Hospital** (ZHENG et al., 2020a): 10.247 exames de ECG com duração de 10 segundos e frequência de amostragem de 500 Hz, todos disponibilizados publicamente.

- **Ningbo First Hospital** (ZHENG et al., 2020b): 34.905 exames de ECG com duração de 10 segundos e frequência de amostragem de 500 Hz, todos disponibilizados publicamente.

Além destes conjuntos de dados, o desafio ainda contou com mais duas fontes não publicamente divulgadas, que constituíam conjuntos de validação e teste do desafio.

Cada exame destes conjuntos de dados está associado a um ou mais diagnósticos, além de informações de sexo e idade do paciente.

Este conjunto de dados é bastante heterogêneo, contendo dados adquiridos na China, Estados Unidos e Alemanha, abrangendo vários grupos étnicos. Além desta heterogeneidade geográfica, cada dataset foi adquirido com certo objetivo e considerando diferentes metodologias de diagnósticos. Deste modo, a proporção de diagnósticos varia entre cada um dos datasets. Por exemplo, no PTB-XL 83,7% dos diagnósticos correspondem a ritmo sinusal normal, enquanto no CPSC2018 apenas 8,9% dos exames possuem este diagnóstico. Esta discrepância também é encontrada no dia a dia dos serviços de saúde, onde hospitais de referência em cardiologia geralmente receberão menos pacientes com ritmo sinusal do que postos de saúde, ou considerando um mesmo hospital, a proporção de pacientes com anormalidades cardíacas será maior na emergência do que no ambulatório.

A Tabela 2 exibe o total de exames para cada classe de interesse. Apesar de cada dataset individual apresentar grande desbalanceamento de classes, ao agrupar todos os datasets, ocorre um melhor balanceamento.

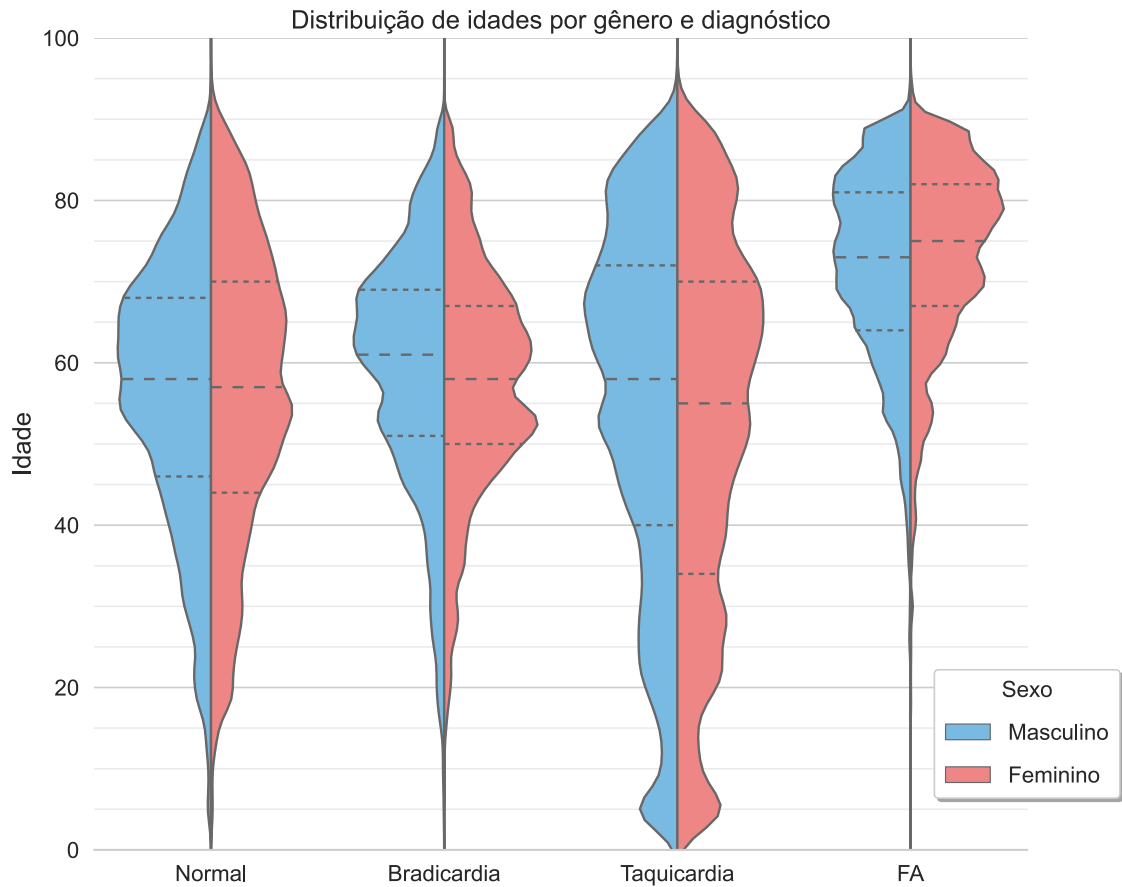
Tabela 2: Distribuição dos exames por classes de ritmos de interesse.

	<b>Ritmo</b>	<b>Nº de exames</b>
RSN	Ritmo Sinusal Normal	28 314
TS	Taquicardia Sinusal	9 624
BS	Bradicardia Sinusal	18 895
FA	Fibrilação/Flutter Atrial	13 597
	Outros ritmos	17 823

Fonte: De autoria própria.

A Figura 31 exibe a distribuição das idades dos pacientes em função do sexo e ritmo diagnóstico. Nota-se uma discrepância principalmente para o ritmo FA, sendo o único que possui mediana de idade superior ao percentil 75 de todos os demais ritmos. Isto é uma consequência da distribuição natural da condição, que possui prevalência maior na população idosa.

Figura 31: Distribuição das idades dos pacientes por gênero e ritmo. Linhas pontilhadas delimitam o intervalo interquartil de cada curva de distribuição.



Fonte: De autoria própria.

Como este dataset foi adquirido utilizando dispositivos clínicos e o objetivo do presente trabalho é desenvolver classificadores de baixo custo computacional que possam ser futuramente utilizados em conjunto com dispositivos vestíveis, foram realizadas adaptações nos sinais deste dataset para simular uma condição minimamente similar àquela encontrada em ECGs vestíveis.

Em geral, dispositivos vestíveis possuem apenas uma ou duas derivações, sendo que no caso de duas derivações é possível aumentar artificialmente para seis derivações segundo o triângulo de Einthoven. A frequência de amostragem destes dispositivos varia e podem atingir características próximas de um ECG clínico, entretanto maiores frequências de amostragem resultam em maior gasto energético. Tendo em consideração estas características, foram consideradas apenas as derivações I e II, ou seja, as derivações III, aVL, aVR e aVF não foram reconstruídas a partir destas duas derivações pois tal reconstrução resultaria em maior gasto computacional, já que os modelos desenvolvidos teriam uma



entrada com seis canais contendo informações redundantes.

Os sinais também foram reamostrados em duas frequências de amostragem típicas de eletrocardiógrafos vestíveis: 200 Hz e 128Hz. Tal operação foi realizada com a técnica de reamostragem com filtragem polifásica, onde o sinal é primeiro sobreamostrado, aplica-se um filtro passa-baixas FIR e em seguida é subamostrado para a frequência desejada (VIRTANEN et al., 2020; VAIDYANATHAN, 1990). O filtro passa-baixas desta técnica atua como filtro antialiasing (SEIDNER, 2005).

Por fim, todos os sinais foram restritos apenas aos seus primeiros 10 segundos, sendo que sinais com duração inferiores foram descartados. A duração de 10 segundos foi escolhida por ser a duração mais comum ao longo do dataset.

## 3.2 Pré-processamentos

Um dos desafios encontrados no processamento de sinais provenientes de dispositivos vestíveis é o elevado nível de ruídos. No caso do ECG clínico, o exame é realizado com o paciente deitado e sob a supervisão de um técnico. Nestas condições ainda é possível ocorrer algum artefato (eletrodo desconectado, movimento do paciente, mal posicionamento dos eletrodos, etc.), mas o técnico, ou posteriormente o médico, podem verificar a qualidade do exame e solicitar um novo em casos de artefatos que comprometam o diagnóstico. Já no contexto da classificação automática em tempo real de sinais de ECG vestível, é necessário que o algoritmo de classificação esteja preparado para lidar com situações de elevado ruído. Alguns dispositivos já possuem estratégias de minimização de artefatos, como posicionamento de eletrodos em posições com menor movimentação, restauração rápida de linha de base, sistemas de detecção de eletrodos desconectados, filtragens mais robustas ou presença de sensores de movimento (que permitem, por exemplo, uma filtragem adaptativa). Com exceção de filtragem mais robusta, as demais abordagens são implementáveis apenas em hardware (a recuperação da linha de base também poderia ser implementada em software, mas é limitada caso ocorra saturação do sinal). Após a aquisição do sinal, em software é possível abordar os problemas de ruído de duas formas: filtragem robusta e/ou análise de qualidade de sinal.

Neste trabalho foi adotado um critério simples de uma análise de qualidade que exclui sinais com trechos de mais de 500 ms de sinal constante, já que foram utilizados sinais de um dataset de ECG clínico. Uma futura aplicação em dispositivos vestíveis necessitaria de mais critérios de qualidade, seja para exclusão de trechos ruidosos ou para dar menor

importância para tais trechos (SATIJA et al., 2018; SHI et al., 2019; RAHMAN et al., 2022).

Em um exame de 12 derivações, caso um artefato ocorra em alguma derivação é possível que a informação perdida seja recuperada a partir das demais derivações. Já no caso de apenas duas derivações, esta possibilidade é reduzida, principalmente se o artefato tiver origem no eletrodo do braço direito. Diante desta dificuldade de recuperação de informações, foi proposto um processamento robusto e que tem como objetivo fornecer redundância de algumas informações mesmo em caso de presença de artefatos.

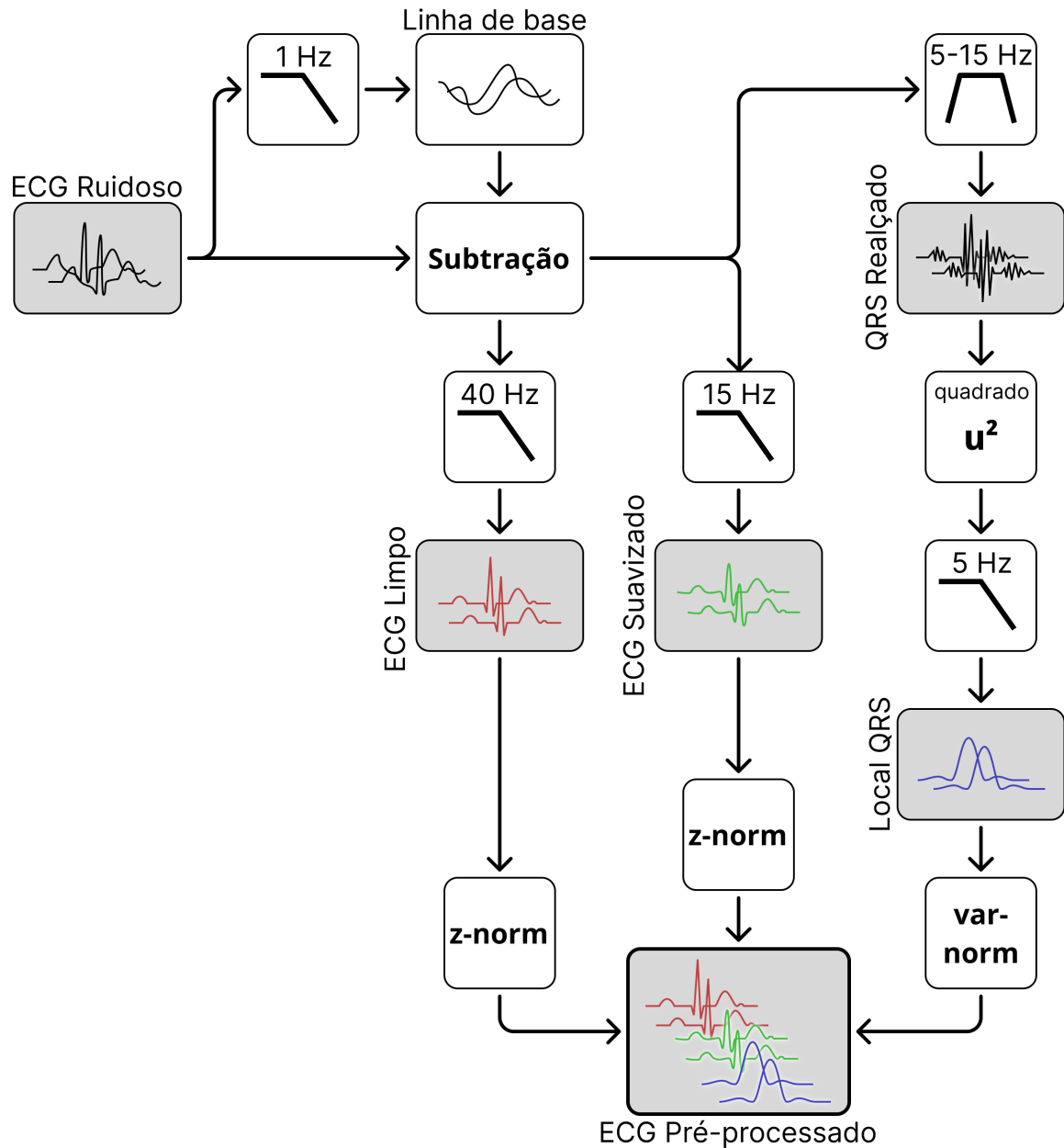
O pré-processamento proposto consiste em estimar a linha de base por meio de um filtro passa-baixas Butterworth de 2<sup>a</sup> ordem<sup>1</sup> em 1 Hz e subtraí-la do sinal. Posteriormente são gerados três sinais: o primeiro consiste de aplicar um filtro passa-baixas Butterworth de 2<sup>a</sup> ordem<sup>1</sup> em 40 Hz; no segundo é aplicado um filtro passa-baixas de 15 Hz; já o terceiro consiste de aplicar um filtro passa-faixa Butterworth de 2<sup>a</sup> ordem<sup>1</sup> de 5 a 15 Hz, elevar o sinal resultante ao quadrado e aplicar um filtro passa-baixas Bessel de 2<sup>a</sup> ordem<sup>1</sup> em 5 Hz. Os primeiros dois sinais são normalizados para variância unitária e média zero, já o terceiro sinal é normalizado apenas para variância unitária. Estes três sinais, de dois canais cada, são então concatenados e formam um sinal com seis canais.

A Figura 32 esquematiza o pré-processamento adotado e a Figura 33 mostra o exemplo de um sinal que passou por tal processamento. O pré-processamento típico para classificação de ECG consiste apenas da remoção de linha de base, filtro passa-baixas com frequência de corte em torno de 40 Hz e normalização (Figura 33-B). Os demais pré-processamentos aplicados buscam gerar uma redundância de parte das informações, mesmo em condições de artefatos e alto ruído. O uso do filtro passa-baixas em 15 Hz (Figura 33-C) atenua parte do complexo QRS do sinal e gera distorções na morfologia do sinal, mas garante que artefatos de alta frequência (como a atividade elétrica muscular não cardíaca) sejam mais atenuados do que no pré-processamento típico, além de destacar as ondas P e T. Por fim, o terceiro sinal utilizado tem semelhanças com os pré-processamentos para detecção de complexos QRS e cálculo de frequência cardíaca (Figura 33-D). O filtro passa-faixa entre 5 Hz e 15 Hz maximiza a energia do complexo QRS (PAN; TOMPKINS, 1985). Ao elevar o sinal resultante ao quadrado, as regiões do QRS ficam mais destacadas e o sinal é retificado. Por fim, o último filtro passa-baixas de 5 Hz atua como um detector de envelope, sendo que o filtro Bessel foi escolhido para esta tarefa pois apresenta menor *overshoot* se comparado aos demais filtros IIR.

---

<sup>1</sup>Todos os filtros foram aplicados em ambos os sentidos do sinal, utilizando a função `scipy.signal.filtfilt`, o que efetivamente dobra a ordem dos filtros.

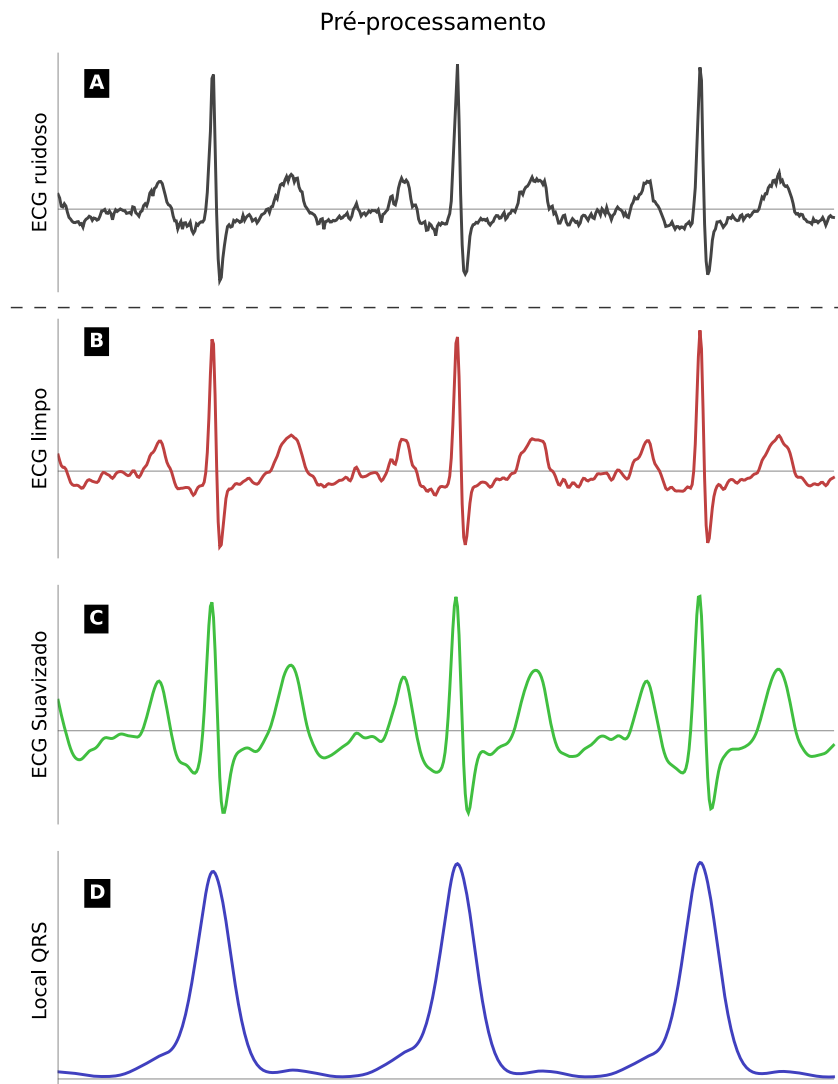
Figura 32: Diagrama de blocos do pré-processamento adotado.



Fonte: De autoria própria.

Nenhum processamento adicional é realizado, sendo que a entrada para os modelos de classificação é composta por um vetor de tamanho  $10 \times frequencia\_amostragem$  e com 6 canais. Também foram realizados experimentos com apenas cada um dos três sinais resultantes deste pré-processamento, para avaliar os impactos de seu uso.

Figura 33: (A) Sinal original de ECG e sinais pré-processados: (B) ECG limpo, (C) sinal suavizado e (D) posição dos complexos QRS.



Fonte: De autoria própria.

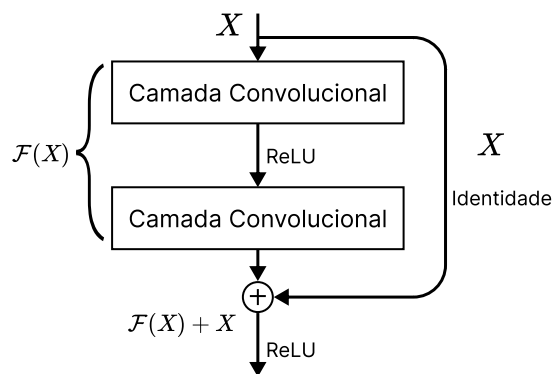
### 3.3 Modelos de aprendizado profundo

Por muito tempo o limitante para a popularização de redes neurais profundas foi o custo computacional. O aumento no poder computacional e a popularização de GPUs viabilizou o seu uso em diversas áreas. O tamanho e complexidade dos modelos de redes neurais acompanhou a disponibilidade de poder computacional e modelos complexos passaram a ser o padrão para a maioria das aplicações, sendo que, por exemplo, modelos com dezenas ou centenas de milhões de parâmetros são comuns em tarefas de visão computacional e modelos ainda maiores, como o GPT-3 (BROWN et al., 2020), possuem centenas de bilhões de parâmetros. Esta tendência de uso de modelos mais complexos pode ser

vista nos trabalhos que concorreram ao desafio de origem do dataset utilizado. Entretanto, algumas tarefas ainda necessitam de modelos mais eficientes, pois são executados em plataformas móveis ou microcontroladores.

No *Physionet Challenge 2021*, a maioria dos times participantes utilizaram técnicas de aprendizado profundo, incluindo redes convolucionais, em especial redes convolucionais residuais (ResNet) (REYNA et al., 2022). As redes residuais possuem blocos residuais (Figura 34), que são blocos especiais onde a saída de um conjunto de camadas convolucionais é somada à entrada do próprio bloco, desse modo permitindo um caminho livre para a propagação dos gradientes no treinamento e possibilitando o treinamento de redes mais profundas (com até mais de mil camadas) (HE et al., 2016). Por ser uma rede muito popular, a mesma (adaptada para sinais unidimensionais) foi escolhida como um dos *baselines* deste trabalho.

Figura 34: Bloco residual genérico.



Fonte: Adaptado de He et al. (2016).

Redes convolucionais genéricas por vezes se assemelham muito com as redes VGGNet (SIMONYAN; ZISSERMAN, 2015), sendo constituídas por sequências de blocos convolucionais intercalados por camadas de *pooling*, uma camada de *flattening* e um bloco de camadas densas. A rede VGG-11 (adaptada para sinais unidimensionais) foi escolhida como um segundo *baseline* para este trabalho.

Katsaouni et al. (2022) desenvolveu uma arquitetura de modelos leve e energeticamente eficiente para a classificação de FA de forma embarcada no ECG vestível. Este arquitetura utiliza uma rede convolucional simples, com apenas três camadas convolucionais, com 1, 2 e 2 filtros por camada, respectivamente, e configurações variáveis de tamanho dos filtros. O bloco denso desta rede possui apenas uma camada com uma unidade e realiza a classificação binária entre FA e não-FA. Esta rede convolucional sozinha

não é capaz de obter bom desempenho, porém é muito eficiente do ponto de vista energético. Para melhorar as métricas do modelo, a estratégia proposta conta com uma etapa recursiva. Cada trecho de sete segundos de sinal passa pela rede convolucional e sua saída é armazenada em uma janela de até 16 posições, sendo que uma função recursiva é aplicada a estes trechos e uma tomada de decisão é realizada a cada recursão, para uma das três possibilidades: FA, não-FA ou precisa de mais uma janela. Essa recursão é realizada até ao máximo das 16 posições, sendo que uma decisão precoce resulta em economia de energia. O modelo como um todo possui pouco mais de 100 parâmetros e apresentou uma acurácia acima de 95% no dataset utilizado para teste. A desvantagem deste modelo é a necessidade de sinais de longa duração e o atraso entre o início de um evento e a sua detecção. Devido à curta duração dos sinais do dataset do *PhysioNet Challenge 2021*, apenas a rede convolucional desta arquitetura foi utilizada para comparação. Foram consideradas as cinco configurações de tamanho dos filtros que, conforme Katsaouni et al. (2022), resultaram em maior eficiência energética.

Os modelos de redes neurais convolucionais 1D propostos nos trabalhos de Hsieh et al. (2020), Fujita e Cimr (2019) e Nurmaini et al. (2020) também foram incluídos como *baselines* de comparação. Apenas a arquitetura de rede neural convolucional destes trabalhos foi reproduzida, ou seja, pré e pós-processamentos originais foram descartados e o pré-processamento aqui proposto foi utilizado. Portanto, a performance aqui obtida não reflete a performance descrita nos artigos de tais modelos.

### 3.3.1 Modelos propostos

Três novas arquiteturas são apresentadas neste trabalho:

- **LiteVGG**: Uma adaptação da arquitetura VGGNet focada na redução do número de parâmetros e custo computacional por meio do uso de camadas *depth-wise* separáveis, menor número de filtros por camada e *global max pooling* como *flattening*;
- **LiteResNet**: Uma adaptação da arquitetura ResNet mais focada na redução do custo computacional por meio do uso de camadas *depth-wise* separáveis e menor número de filtros por camada;
- **DwNet**: Uma nova arquitetura que utiliza apenas camadas de convolução *depthwise*, ou seja, apenas a primeira etapa da convolução separável utilizada nas arquiteturas acima. Um número reduzido de filtros e *strides* grandes são utilizados para manter o número de parâmetros e o custo computacional baixos.

As duas primeiras arquiteturas são adaptações de modelos existentes. A metodologia adotada para estas adaptações consiste de, quando possível: 1) substituir todas camadas convolucionais tradicionais por convoluções *depth-wise* separáveis, 2) reduzir o número de filtros em cada convolução, 3) adotar *poolings* globais com estratégia de achatamento e 4) reduzir o número de unidades nas camadas densas. A princípio, esta metodologia pode ser utilizada, mesmo que parcialmente, em todas arquiteturas de redes neurais convolucionais tradicionais que possuam alto custo computacional. Entretanto, como há uma alteração substancial na arquitetura, testes sempre são necessários para avaliar os impactos nas capacidades de classificação.

### 3.3.1.1 LiteVGG

A arquitetura VGGNet consiste basicamente de cinco blocos convolucionais seguidos por uma camada de achatamento que conecta a um bloco de três camadas densas. Os blocos convolucionais podem ser compostos por diferentes números de camadas convolucionais, sendo que a Tabela 3 apresenta as configurações mais comuns. Uma camada de *max pooling* é adicionada após cada bloco convolucional.

A implementação utilizada neste trabalho ainda adiciona camadas de normalização em lotes antes de cada bloco convolucional, objetivando uma maior velocidade de treinamento, e camadas de *dropout* espacial após cada camada de *max pooling*, como regularização.

A maior parte das operações de ponto flutuante necessárias para inferência de uma rede VGG estão nas camadas convolucionais, já a maior parte dos parâmetros estão nas camadas densas. Portanto, foi necessário adaptar a arquitetura tendo em vista ambas otimizações, pois um modelo a ser executado em plataformas móveis geralmente enfrenta limitações tanto de processamento quanto de armazenamento.

Na arquitetura LiteVGG, a redução de custo computacional, em termos de número de operações, é obtida ao utilizar camadas convolucionais *depthwise* separáveis, com menor número de filtros. Em uma rede VGG tradicional para imagens, o número de filtros das camadas convolucionais é dobrado (com exceção do último bloco convolucional) após cada camada de *max pooling*. Deste modo, o mapa de ativação tem o seu tamanho reduzido pela metade, já que o *pooling* reduz largura e altura do mapa pela metade, mas a profundidade (canais) é dobrada. Com a redução do tamanho do mapa de ativação, ocorre uma redução do custo computacional de cada bloco em relação ao anterior. Considerando esta estratégia em uma rede unidimensional, dobrar o número de filtros após um *pooling* resultaria em manter o tamanho do mapa de ativação. Assim, o custo computacional passa

a ser praticamente constante em todos os blocos. A abordagem utilizada na arquitetura LiteVGG consiste de incrementar o número de filtros após o *pooling* por um fator de  $\sqrt{2}$ , o que resulta na redução do mapa de ativação por este mesmo fator. Uma última adaptação dos blocos convolucionais consiste de reduzir o número de filtros do primeiro bloco de 64 para apenas 16 filtros.

A redução do número de parâmetros e, portanto, do tamanho dos modelos foi alcançado principalmente ao utilizar uma camada de *max pooling* global como estratégia de achatamento. O número de unidades nas camadas densas também foi reduzido de 4096 e 4096 para 64 e 32, sendo que a última camada densa possui número de unidades compatível com o número de classes.

As adaptações realizadas com foco na redução do número de parâmetros também reduziram o número de operações e vice versa. A Figura 35 apresenta uma comparação visual entre os modelos LiteVGG-11 e VGG-11.

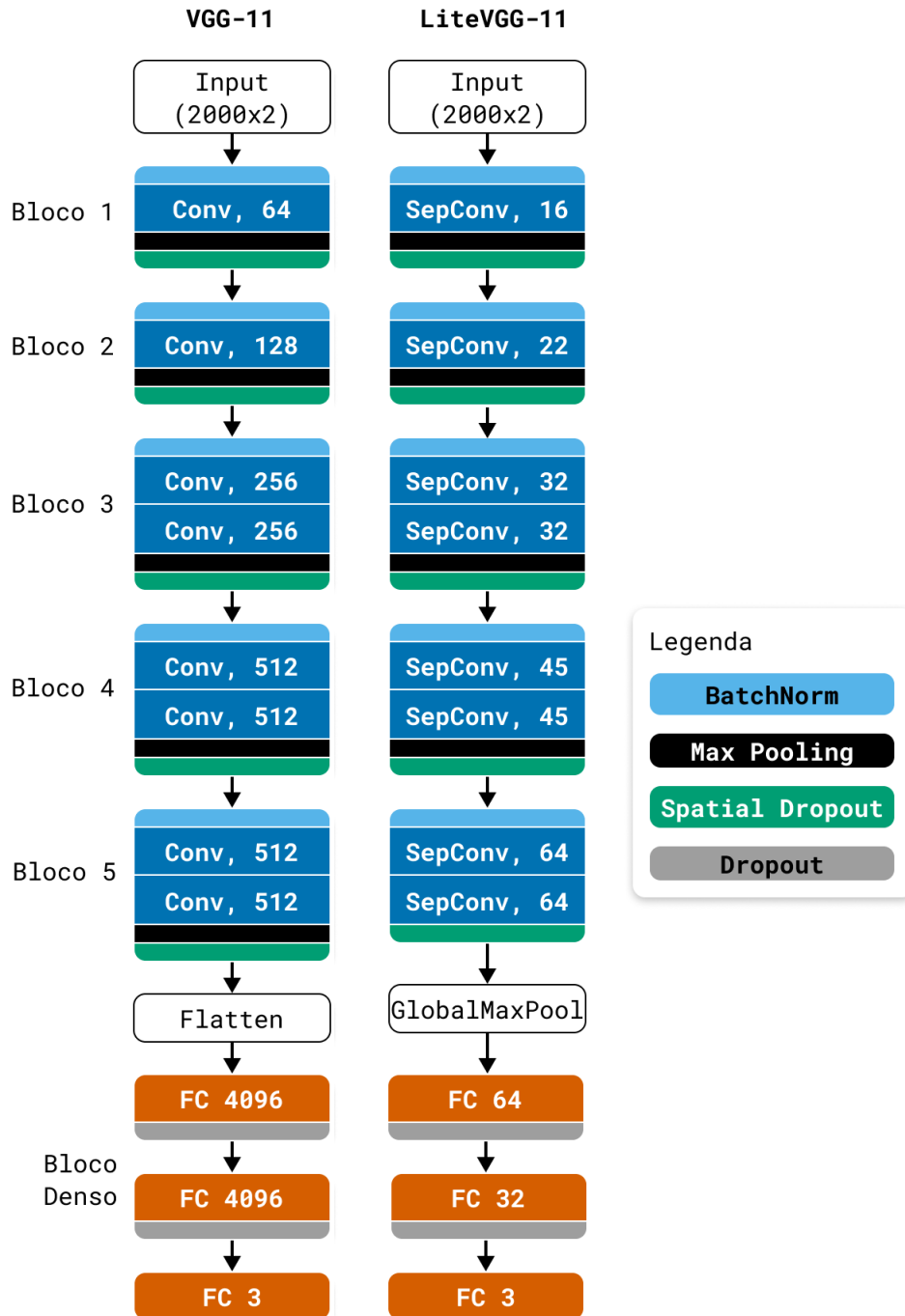
Tabela 3: Configurações convencionais de camadas convolucionais em cada bloco da arquitetura VGGNet. Camadas convolucionais estão nomeadas como *conv $x$ - $y$* , sendo  $x$  o tamanho do kernel e  $y$  o número de filtros da camada.

<b>Blocos</b>	<b>VGG-11</b>	<b>VGG-13</b>	<b>VGG-16</b>	<b>VGG-19</b>
<i>Conv 1</i>	conv3-64	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
<i>Conv 2</i>	conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
<i>Conv 3</i>	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 conv3-256	conv3-256 conv3-256 conv3-256
<i>Conv 4</i>	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512
<i>Conv 5</i>	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512
<i>Denso</i>			FC-4096 FC-4096 FC-out	

Fonte: Adaptado de (SIMONYAN; ZISSERMAN, 2015).



Figura 35: Comparação entre a arquitetura dos modelos VGG-11 e LiteVGG-11.



Fonte: De autoria própria.

A Tabela 4 mostra um passo a passo de como as adaptações realizadas na arquitetura VGG-11 para se transformar na arquitetura LiteVGG-11 impactam nos números de parâmetros e de operações de ponto flutuante (FLOP) necessárias para realizar uma inferência, considerando quatro classes e uma entrada de tamanho  $2000 \times 6$ . A codificação dos modelos intermediários apresentada na tabela indica:

- A. Utilização de convoluções *depthwise* separável;
- B. Incremento por um fator de  $\sqrt{2}$  no número de filtros após cada *pooling*;
- C. Uso de um menor número de filtros;
- D. *Pooling* global como achatamento;

Observa-se que a redução de quase três ordens de grandeza no número de operações para inferência e uma redução de quase cinco ordens de grandeza no número de parâmetros.

Tabela 4: Comparação entre o número de operações de ponto flutuante (FLOP) e parâmetros dos modelos VGG-11 e LiteVGG-11, bem como das adaptações intermediárias. Linhas indicadas por uma seta para baixo ( $\downarrow$ ) indicam a redução percentual entre o modelo anterior e o posterior.

	Blocos				Total	
	Convolucionais		Densos		FLOP	# Param
	FLOP	# Param	FLOP	# Param		
<b>VGG-11</b>	1,48 G	3,14 M	297 M	148 M	1,78 G	152 M
$\downarrow$	66,21%	66,29%	0,00%	0,00%	55,13%	1,37%
<b>A</b>	500 M	1,05 M	297 M	148 M	798 M	149 M
$\downarrow$	73,05%	79,10%	44,36%	44,36%	62,35%	44,60%
<b>B</b>	134 M	221 k	165 M	82 M	300 M	83 M
$\downarrow$	92,67%	93,08%	59,80%	59,79%	74,55%	59,88%
<b>C</b>	9,89 M	15311	66 M	33 M	76 M	33 M
$\downarrow$	0,00%	0,00%	48,78%	48,78%	42,48%	48,76%
<b>D</b>	9,89 M	15311	34 M	17 M	44 M	17 M
$\downarrow$	0,00%	0,00%	99,95%	99,96%	77,49%	99,87%
<b>LiteVGG-11</b>	9,89 M	15311	16612	6372	9,90 M	21683
<b>Redução Total</b>	99,33%	99,51%	99,99%	99,996%	99,44%	99,99%

Fonte: De autoria própria.

### 3.3.1.2 LiteResNet

Esta arquitetura é idêntica à ResNet original, a menos do uso de convoluções *depthwise* separáveis unidimensionais e do número reduzido de filtros em cada módulo de blocos residuais.

Os modelos iniciam com uma convolução com kernel de tamanho 7 e *stride* 2, sendo seguido por uma camada de normalização por lote e de um *max pooling* de *stride* 2 e pool de tamanho 3. Em seguida são adicionados quatro módulos de blocos residuais. As convoluções de cada módulo possuem, respectivamente, 16, 22, 32 e 45 filtros, similar aos blocos do LiteVGG. Os módulos são compostos por sequências de blocos convolucionais e o número de blocos em cada módulo determina a profundidade da rede. Neste trabalho foram testados modelos com profundidade 18 (LiteResNet-18), que possui dois blocos em cada módulo residual, e modelos com profundidade 34 (LiteResNet-34), que possui módulos com 3, 4, 6 e 3 blocos, respectivamente.

Os blocos residuais são montados conforme a Figura 34, com a adição de uma camada de normalização por lotes após cada convolução. Com exceção do primeiro módulo, os demais iniciam com uma redução no tamanho do mapa de ativação. Tal redução é obtida por meio de uma convolução com *stride* de tamanho 2. Neste caso, a função identidade é substituída por uma convolução com kernel de tamanho 1 e com *stride* de tamanho 2, de modo a manter a dimensão dos mapas de ativação somados.

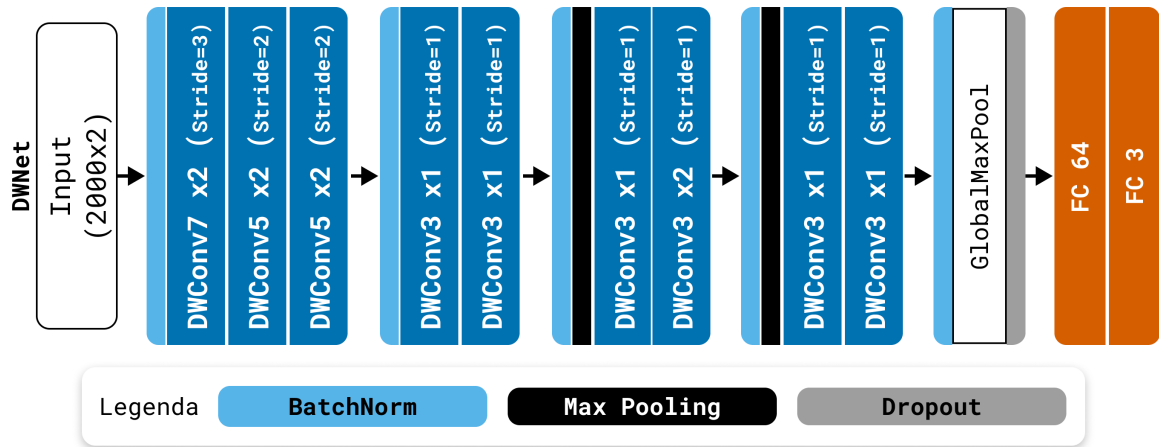
O Apêndice B apresenta diagramas da arquitetura dos modelos LiteResNet-18 e LiteResNet-34, bem como suas contrapartes na arquitetura ResNet tradicional.

### 3.3.1.3 DwNet

Esta arquitetura foi proposta como uma forma de redução ainda maior na complexidade dos modelos. Nesta arquitetura, as camadas convolucionais são do tipo *depthwise*, ou seja, as convoluções são aplicadas com kernels distintos para cada canal, sendo que é possível multiplicar o número de canais por um fator  $dm$ . Uma desvantagem desta camada é que as informações de canais distintos nunca são combinados, mas com a vantagem computacional de uma operação de convolução a menos.

O modelo proposto é apresentado na Figura 36, sendo que cada camada convolucional é denotada como DWConv $\alpha$  X $\beta$ , sendo  $\alpha$  o tamanho do kernel e  $\beta$  o fator de multiplicação no número de canais.

Figura 36: Arquitetura do modelo DWNet.



Fonte: De autoria própria.

### 3.4 Métricas de custo computacional

O custo computacional dos modelos foi avaliado em termos de métricas teóricas e empíricas. As métricas teóricas avaliadas dependem exclusivamente do modelo, são elas:

- número total de parâmetros (treináveis e não treináveis): reflete indiretamente o tamanho ocupado pelos pesos do modelo em disco; e
- número de operações de ponto flutuante (FLOP): reflete indiretamente o tempo necessário para realizar uma inferência;

Entretanto, um modelo com maior número de FLOP pode realizar uma inferência em menor tempo do que um modelo de menor número de FLOP a depender da existência de instruções otimizadas para determinadas camadas em alguns hardwares. Desse modo, também foi proposto utilizar as métricas de tempo de inferência em diferentes plataformas de hardware, além do *footprint* de memória RAM durante uma inferência.

Considerando uma possibilidade de embarcar o modelo em microcontroladores, também foram estimados os requisitos mínimos para execução dos modelos em microcontroladores ARM, em termos de memórias FLASH e RAM. O tempo de inferência nestas condições não foi estimado pois a implementação dos modelos em hardware está além do escopo deste trabalho.

### 3.5 Métricas de classificação

O desempenho de um modelo de aprendizado de máquina é mensurado ao calcular algumas métricas em um subconjunto dos dados que não foi utilizado para treinamento. Diferentes tarefas de aprendizado de máquina possuem diferentes métricas de avaliação de modelos.

Considerando a classificação binária, os dados de teste, de igual modo aos de treinamento, possuem rótulos que os identificam como pertencente à classe positiva ou à classe negativa e ao aplicar os dados no modelo, o mesmo pode concordar ou discordar das classes conhecidas. Nesta situação existem quatro possíveis combinações de classe predita e classe verdadeira:

- Verdadeiro Positivo (VP): modelo acertou ao classificar a amostra como pertencente à classe positiva;
- Verdadeiro Negativo (VN): modelo acertou ao classificar a amostra como pertencente à classe negativa;
- Falso Positivo (FP): modelo classificou como pertencente à classe positiva, mas a amostra pertence à classe negativa; e
- Falso Negativo (FN): modelo classificou como pertencente à classe negativa, mas a amostra pertence à classe positiva.

O conjunto de dados de teste é apresentado ao modelo classificador e a frequência de cada uma destas combinações de classe predita e verdadeira é calculada. Estas frequências geralmente são organizadas em uma tabela conhecida como matriz de confusão, conforme esquema exibido na Figura 37, onde também são apresentados o total de amostras por classe predita e por classe verdadeira.

Muitas métricas de classificação são derivadas da matriz de confusão, sendo que as mais comumente utilizadas para comparação de modelos de aprendizado de máquina são: acurácia (Acc), sensibilidade (TPR, de *True Positive Rate*) (também conhecido como *recall*), especificidade (TNR, de *True Negative Rate*), valor preditivo positivo (PPV) (também conhecido como precisão), valor preditivo negativo (NPV) e F1-score. A seguir são listadas como são derivadas cada uma destas métricas:

$$Acc = \frac{VP + VF}{VP + VF + FP + FN} \quad (3.1)$$

Figura 37: Matriz de confusão.

		Classe Predita	
		PP	PN
Classe Verdadeira	P	VP	FN
	N	FP	VN

Fonte: De autoria própria.

$$TPR = \frac{VP}{P} = \frac{VP}{VP + FN} \quad (3.2)$$

$$TNR = \frac{VN}{N} = \frac{VN}{VN + FP} \quad (3.3)$$

$$PPV = \frac{VP}{PP} = \frac{VP}{VP + FP} \quad (3.4)$$

$$NPV = \frac{VN}{PN} = \frac{VN}{VN + FN} \quad (3.5)$$

$$F1_{score} = 2 \frac{PPV \times SE}{PPV + SE} = \frac{2VP}{2VP + FP + FN} \quad (3.6)$$

Cada métrica avalia diferentes capacidades do modelo. A acurácia, Equação 3.1, mensura a capacidade do modelo acertar a predição, mas não considera qualquer viés que o modelo possa ter. A sensibilidade, Equação 3.2, mensura a capacidade do modelo identificar a classe positiva, enquanto a especificidade, Equação 3.3, mensura a capacidade do modelo identificar a classe negativa. Uma relação de compromisso existe entre a sensibilidade e especificidade para um classificador não perfeito, sendo que geralmente é possível ajustar o modelo de modo a melhorar uma destas métricas ao custo de diminuir a outra. O PPV, Equação 3.4, mensura quão confiável é uma predição que aponta para a classe positiva, enquanto o NPV, Equação 3.5, mensura o quão confiável é uma predição da classe negativa. Uma relação de compromisso também existe entre o PPV e o NPV, onde é possível melhorar uma métrica em detrimento da outra. Por fim, o F1-score, Equação 3.6, é uma média harmônica entre a sensibilidade e o PPV, medindo ao mesmo

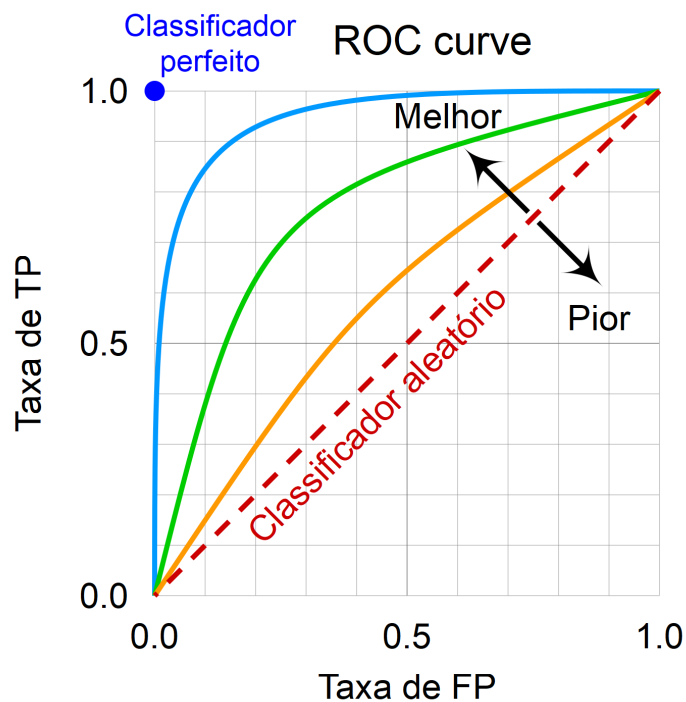
tempo o quão bom o modelo é em identificar uma amostra da classe verdadeira e o quão confiável é uma predição referente à classe verdadeira.

Considerando um classificador com saída contínua ( $y_{out}$ ), como as redes neurais, um limiar  $y_{thresh}$  deve ser utilizado para determinar uma saída binária. Nesta situação, a curva ROC é uma curva obtida ao traçar a taxa de verdadeiros positivos (ou sensibilidade) em função da taxa de falsos positivos ( $1 - TNR$ ) conforme varia-se o limiar de decisão. A Figura 38 mostra um exemplo de diversas curvas ROC e mostra que quanto mais fidedigna a classificação modelo, maior será a área delimitada pela curva, sendo que um classificador perfeito possui uma área igual a 1 e um classificador aleatório possui área igual a 0,5. Esta área abaixo da curva ROC (AUC) é uma métrica que indica quão bom o modelo é em gerar uma saída na qual é possível separar as classes.

Para uma classificação com mais de duas classes, é possível considerar uma classe como verdadeira enquanto as demais são falsas e calcular as métricas individuais de cada classe. As métricas globais do modelo são definidas como a média entre as métricas de cada classe. Neste trabalho, foi utilizada uma média ponderada pela frequência de ocorrência de cada uma classes.

Neste trabalho serão apresentadas todas estas métricas e uma métrica adicional deri-

Figura 38: Curva ROC.



Fonte: Adaptado de Cmglee e Thoma (2021).

vada da sensibilidade, a taxa de falsos negativos ( $FNR = 1 - TPR$ ). Esta última auxilia a racionalizar as diferenças quando as sensibilidades dos modelos comparados são muito altas. Os modelos, no entanto, serão comparados prioritariamente em termos de TNR, F1-score e AUC.

Ao considerar apenas a classe FA, a métrica TNR é de grande importância. Isto ocorre pois, durante o monitoramento contínuo de ritmo, é desejado que seja reduzido o número de eventos de FA não detectados e, de certa forma, é tolerável que o modelo identifique tais eventos em um sinal normal. Esta é uma situação muito comum em modelos de diagnósticos de doenças, pois nestes casos um falso negativo muitas vezes não será investigado, enquanto um falso positivo pode ser descartado por testes adicionais.

### 3.6 Experimentos

A classificação de FA foi abordada de três formas distintas: binária, multiclasse e multilabel. Em cada uma destas abordagens, foram realizados testes preliminares com todos os modelos e um menor volume de dados. Nestes testes preliminares foram incluídas as variantes dos modelos VGG e LiteVGG com 11 e 19 camadas, e as variantes dos modelos ResNet e LiteResNet com profundidades 18 e 34, além da DWNet e demais modelos *baseline* da literatura. Estes testes foram importante para descartar modelos com desempenho muito degradado e para escolher os melhores modelos LiteVGG e LiteResNet, reduzindo o custo computacional da comparação mais robusta e com um maior volume de dados. Em ambos testes, os modelos foram comparados quanto às métricas obtidas por validação cruzada.

Como mostrado na Tabela 2, as classes diagnósticas são bastante desbalanceadas, o que dificulta o treinamento de modelos de aprendizado profundo e pode introduzir viés para as classes mais prevalentes. Para reduzir o desbalanceamento das classes, foi utilizada uma validação cruzada com amostragem de Monte Carlo estratificada por classes. Deste modo, em cada repetição da validação um determinado número de amostras aleatórias pertencentes a cada classe foi selecionada para treinamento, validação e teste. Nos testes preliminares foram utilizadas 10 validações, enquanto nos testes robustos foram utilizadas 20 validações.

Os modelos foram implementados e treinados em Tensorflow/Keras. O treinamento foi realizado com o otimizador Adam (KINGMA; BA, 2015), utilizando entropia cruzada (binária/catórica) como perda a ser otimizada. O tamanho dos lotes de treinamento



foi definido de forma heurística em 128, de modo a obter o máximo de utilização das GPUs disponíveis para treinamento sem ocorrer estouro da memória em nenhum dos modelos testados. A taxa de aprendizado inicial foi determinada por meio da técnica proposta por Smith (2015) aplicada no modelo ResNet-34, onde observou-se que taxas de aprendizado próximas a  $10^{-3}$  conduziam à uma maior redução no valor da função de perda nas primeiras épocas do treinamento. Portanto, a taxa de aprendizado inicial foi configurada como  $10^{-3}$ , sendo aplicado um fator de redução em 25% após cada platô de 25 épocas sem redução da função de perda calculada nos dados de validação. O critério de parada do treinamento foi definido como 200 épocas ou um platô de 40 épocas sem redução da perda nos dados de validação.

### 3.6.1 Classificação Binária

Este experimento possui como objetivo apenas a detecção de FA em detrimento aos demais possíveis diagnósticos. Todos os exames foram estratificados como FA ou não-FA. Deste modo existe um grande desbalanceamento entre as classes, visto que o dataset possui 13.597 exames de FA, o que equivale a 15,4% do total. Nos testes preliminares foram amostrados 1024 exames de cada classe para treinamento, 128 para validação e 128 para teste. Já nos testes robustos, estes números foram incrementados para 7168, 1024 e 1024, respectivamente.

### 3.6.2 Classificação Multiclasse

A detecção apenas de FA já possibilitaria, por exemplo, o monitoramento de ritmo de um paciente no pós-operatório de uma cirurgia cardíaca, visto que esta é a arritmia mais comum neste período. Entretanto, ainda existe a possibilidade de que o paciente desenvolva algum outro ritmo anormal. Este experimento agregou os diagnósticos presentes no dataset em três classes:

1. *RSN*: exames com diagnósticos exclusivamente de ritmos sinusais normais. Bradicardia e taquicardia foram considerados ritmos normais nesta análise, pois podem ocorrer em situações fisiológicas comuns (por exemplo, bradicardia ao dormir ou taquicardia ao fazer um esforço físico);
2. *FA*: exames com diagnósticos de fibrilação ou flutter atrial, independente de haver outros diagnósticos conjuntos;

3. *Outros*: exames que não pertencem às classes anteriores, podendo inclusive conter ritmos sinusais com alguma alteração concomitante.

Estas três classes foram pensadas para funcionar como três níveis de alerta para um possível sistema de monitoramento remoto dos pacientes. A classe *RSN* indicaria que o paciente está com ritmo cardíaco dentro dos padrões. A classe *outros* seria um nível de alerta baixo, onde poderia ser solicitada uma avaliação do ritmo por um cardiologista, mas sem muita prioridade, já que poderia ser um ritmo anormal típico do pacientes (como bloqueios, extrassístole, entre outros). Já a classe *FA* indicaria um nível de alerta grave.

Além de fornecer mais um nível de alerta, esta estratégia também reduz o desbalanceamento original existente entre as classes, que nesta estratificação passa a ser de: 13.597 exames *FA*, 28.909 exames *RSN* e 45.747 exames da classe *outros*. O que gera, respectivamente, uma proporção de 15,4% / 32,8% / 51,8%, permitindo um teste robusto com um maior número de exames por classe. Nos testes preliminares manteve-se as quantidades utilizadas no experimento binário, ou seja, 1024, 128 e 128 exames por classe para treino, validação e teste. O total de exames utilizados, no entanto, aumentou em 50%, já que agora existem um total de três classes. Nos testes robustos, optou-se por utilizar 10.240 exames por classe para treino, 1024 para validação e 1024 para teste.

### 3.6.3 Classificação Multilabel

Por fim, uma classificação multilabel foi proposta. Diferente da classificação multiclasse, aqui é possível que uma amostra pertença a mais de uma classe simultaneamente.

Este experimento foi planejado para verificar se a divisão da classe *RSN* do teste multiclasse em bradicardia, normal e taquicardia resulta em melhor desempenho dos modelos. Esta suposição se baseia no fato de que o ritmo *FA* pode se apresentar em estado de bradicardia, frequência cardíaca normal ou taquicardia. Desse modo, separar os ritmos normais nestes três regimes de frequência cardíaca poderia gerar modelos classificatórios com melhor distinção da *FA*.

As classes utilizadas neste experimento foram *FA*, *RSN*, *BS* e *TS*. Do ponto de vista de conjuntos, estas quatro classes não deveriam possuir intersecção, porém existem ocorrências de sinais com diagnósticos apontando para mais de uma destas classes simultaneamente. A estratégia multilabel foi adotada para que não fosse necessário tomar decisões sobre os rótulos presentes no dataset, tal como foi feito no experimento multiclasse (onde, por exemplo, um exame com *FA* e *TS* foi considerado *FA*, ou um exame

com RSN e depressão do segmento ST foi considerado *outros*), assim não introduzindo possíveis novos vieses. Outra distinção em relação ao experimento multiclasse está no fato de que ritmos sinusais concomitantes com outras condições agora são considerados ritmos sinusais. Uma correspondência para a classe *outros* seria quando um exame não pertence a nenhuma das quatro classes propostas. O total de exames de cada uma das classes é aquele discriminado na Tabela 2.

O número de exames por classe para os testes preliminares e robustos são os mesmos que foram utilizados no teste de classificação binária. O teste robusto não pode utilizar um número maior de amostras de treinamento porque o valor de 7168, 1024 e 1024 amostras para treino, validação e teste já é muito próximo ao número total de exames pertencentes à classe TS. Foram amostrados exames de cada uma das quatro classes consideradas no experimento, além de amostras de exames que não pertencem a nenhuma destas classes. E, como este é um experimento multilabel, exames que pertencem a mais de uma classe foram contabilizados em todas as suas classes; portanto, o número total de exames em cada subconjunto varia ligeiramente de amostragem para amostragem.

### 3.7 Hardwares utilizados

Os treinamentos dos modelos foram realizados utilizando a estrutura de HPC (computação de alto desempenho) do Instituto do Coração (InCor) do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (HCFMUSP). Para treinamento foram utilizadas máquinas equipadas com processador Intel Xeon Skylake (4x 3.0 GHz), com 90 GB de memória RAM e placa gráfica Tesla V100-SXM2 de 16GB de VRAM.

Os testes de desempenho computacional dos modelos foram realizados em um computador e três smartphones, são eles:

- **Notebook:** Acer E5-571 com processador Intel Core i5-5200U (4x 2,2 GHz), com 12 GB de RAM e sem placa gráfica;
- **XPower:** smartphone LG X power, rodando Android 6, com processador Cortex-A53 MediaTek MT6735 (4x 1,3 GHz) e 2 GB de RAM;
- **MotoG8:** smartphone Motorola Moto G8, rodando Android 11, com processador Snapdragon 665 Qualcomm SDM665 (4x 2,0 GHz + 4x 1,8 GHz) e 4 GB de RAM;
- **Note11:** smartphone Xiaomi Redmi Note 11, rodando Android 11, com processador Snapdragon 680 4G Qualcomm SM6225 (4x 2,4 GHz + 4x 1,9 GHz) e 6 GB de RAM.

Os smartphones de teste incluem modelos de entrada e intermediários, lançados entre 2016 e 2022. Esta gama de aparelhos, embora pequena, permite estimar o *throughput* dos modelos em ambientes com diferentes capacidades computacionais. Todos os testes foram realizados em dispositivos físicos, ou seja, não foram utilizados emuladores de hardware.

## 4 RESULTADOS E DISCUSSÕES

Neste capítulo serão apresentados os custos computacionais obtidos nos diferentes ambientes de teste, bem como os resultados dos experimentos de classificação binária, multi-classe e multilabel. Por fim, será apresentado um estudo do impacto do pré-processamento adotado em relação à utilização do pré-processamento mais simples e tradicional (que já faz parte da abordagem proposta).

### 4.1 Custo computacional

Como descrito na metodologia, os custos computacionais de todos os modelos propostos e utilizados como *baseline* de comparação (LiteResNet, LiteVGG, DWNNet, VGG, ResNet, Katsaouni et al. (2022), Hsieh et al. (2020), Nurmaini et al. (2020) e Fujita e Cimr (2019)) foram mensurados em termos de número de parâmetros e operações de ponto flutuante (FLOP), sendo também estimados o *footprint* de memória dos modelos e o tempo para inferência em diferentes plataformas. Nesta seção, serão apresentados os custos computacionais referentes à frequência de amostragem de 128 Hz (a menos que seja explicitada outro valor), visto que o tamanho da entrada dos modelos aumenta com o incremento da frequência de amostragem, conseqüentemente alterando os custos computacionais.

O número de parâmetros é diretamente proporcional ao tamanho ocupado em disco pelo modelo, sendo um fator limitante apenas em casos de modelos extremamente grandes ou plataformas computacionais mais limitadas, como smartphones antigos e microcontroladores. O número de FLOP indica indiretamente a quantidade de instruções que um processador necessitará para realizar uma inferência. Como algumas plataformas possuem instruções especializadas (ou mesmo hardware dedicado) para lidar com algumas operações de redes neurais, como é o caso das bibliotecas NNAPI<sup>1</sup> e XNNPACK<sup>2</sup>, o número de instruções de hardware executadas pode variar dependendo dos recursos dis-

<sup>1</sup><https://developer.android.com/ndk/guides/neuralnetworks>

<sup>2</sup><https://github.com/google/XNNPACK>

poníveis em cada plataforma. Portanto, um menor número de FLOP nem sempre garante uma inferência mais rápida.

Para avaliar o custo computacional em dispositivos reais, foi utilizada uma ferramenta de *benchmark* oficial do TensorFlow Lite<sup>1</sup>. Os modelos a serem testados foram convertidos para TensorFlow Lite sem nenhum tipo de otimização e foram avaliados pela ferramenta utilizando diferentes tipos de aceleração de hardware. A execução em CPU com 4 threads e instruções XNNPACK se mostrou a configuração com menor tempo de inferência em todos os modelos e hardwares de teste, e os resultados médios obtidos em 100 repetições para esta condição são apresentados na Tabela 5, juntamente com o número de parâmetros e FLOPS.

O *footprint* de memória é uma estimativa feita pela ferramenta de *benchmark* ao calcular a diferença no uso de memória antes e durante uma inferência, podendo apresentar interferências de outros processos sendo executados no dispositivo. Para minimizar esta interferência, os testes foram executados com todos processos de primeiro plano fechados. O valor de memória apresentado na Tabela 5 se refere ao valor médio obtidos dentre os

Tabela 5: Custo computacional de todos os modelos testados (*baselines* e propostos). Asterisco (\*) indica modelos avaliados com frequência de amostragem de 200 Hz.

Modelo	Parâmetros	FLOP	Mémória (MB)	Tempos de inferência (ms)			
				MotoG8	Note11	XPower	Notebook
<b>Katsaoumi2022 5</b>	269	131,27 k	17,13	0,08	0,07	0,20	0,12
<b>Katsaoumi2022 1</b>	281	149,09 k	18,57	0,22	0,07	0,23	1,32
<b>Katsaoumi2022 2</b>	285	166,76 k	17,25	0,28	0,08	0,23	0,13
<b>Katsaoumi2022 3</b>	297	184,44 k	17,87	0,10	0,08	0,24	0,14
<b>Katsaoumi2022 4</b>	309	202,02 k	16,84	0,11	0,08	0,27	0,21
<b>DWNet</b>	4,03 k	355,84 k	21,44	0,91	0,95	1,14	0,98
<b>LiteResNet-18</b>	22,50 k	3,89 M	13,43	6,01	2,27	3,95	2,04
<b>LiteVGG-11</b>	20,57 k	4,60 M	21,75	1,53	1,56	3,15	3,01
<b>Fujita</b>	24,89 k	7,21 M	17,57	0,52	0,41	1,64	0,69
<b>LiteResNet-34</b>	42,32 k	7,31 M	12,85	4,46	2,18	6,77	3,89
<b>LiteVGG-19</b>	36,88 k	10,35 M	20,56	2,41	2,22	5,17	1,88
<b>Hsieh</b>	689,44 k	38,33 M	14,89	2,82	1,68	5,47	25,97
<b>ResNet-18</b>	8,74 M	999,44 M	56,28	37,48	30,07	81,95	23,40
<b>VGG11</b>	103,76 M	1,05 G	852,78	79,72	55,72	167,86	43,28
<b>Nurmaini</b>	23,32 M	1,39 G	193,35	65,03	41,83	127,13	30,03
<b>ResNet-34</b>	16,61 M	2,06 G	91,56	73,23	57,71	197,40	39,06
<b>VGG19</b>	107,37 M	2,15 G	681,50	105,80	81,24	359,14	61,65
<b>DWNet*</b>	4,03 k	567,52 k	12,77	1,22	0,98	5,18	0,91
<b>LiteVGG-11*</b>	20,57 k	7,19 M	13,62	2,57	1,82	4,89	1,75
<b>LiteResNet-34*</b>	42,32 k	11,37 M	14,31	5,58	3,00	17,22	4,84

Fonte: De autoria própria.

<sup>1</sup><https://www.tensorflow.org/lite/performance/measurement>

hardwares do tipo *smartphone* utilizados para teste.

Nota-se que alguns modelos de baixo número de FLOP possuem tempo de inferência maior do que outros modelos de maior número de FLOPS, como observa-se nos modelos LiteResNet-18 e o modelo de Fujita, onde o primeiro possui metade dos FLOPS do segundo mas gasta em média quatro vezes mais tempo para realizar uma inferência. Como mencionado anteriormente, isto ocorre pois algumas operações e camadas possuem implementações mais otimizadas do que outras.

Focando nos modelos propostos neste trabalho, o DWNet é o modelo mais leve, seguido pelo LiteVGG-11 e pelo LiteResNet-34. Em média, o modelo VGG-11 leva 33 vezes mais tempo para realizar uma inferência, gastando 51 vezes mais memória que a sua contraparte proposta, o LiteVGG-11. Já o modelo ResNet-34 gasta 12 vezes mais tempo e 17 vezes mais memória para realizar uma inferência quando comparado ao LiteResNet-34.

De todos os modelos utilizados para comparação, os modelos propostos por Katsaouni et al. (2022) são os que apresentam o menor número de parâmetros, FLOPS e tempo de inferência, seguidos pelo modelo proposto por Fujita e Cimr (2019) e Hsieh et al. (2020), ainda apresentando tempo de inferência inferior a 10 ms nos smartphones. O modelo de Nurmaini et al. (2020) possui custo computacional elevado, se comparando com o VGG-11 e ResNet-34. Como era esperado, os modelos para sinais com frequência de amostragem de 128 Hz apresentaram um menor custo computacional, cerca de 63% de FLOP ao comparar com os mesmos modelos para sinais de 200 Hz.

Todos os modelos, inclusive aqueles pesados, são capazes de gerar uma inferência em menos de 1 segundo, e possibilitariam um monitoramento contínuo de ritmo a esta taxa. Entretanto, tal afirmação pode não ser verdadeira em dispositivos ainda mais restritos do que o XPower, ou em situações em que o smartphone possui outras aplicações em aberto. Além disso, uma inferência em menor tempo indica um menor gasto energético, que é um fator a ser considerado no desenvolvimento de aplicativos móveis ou de softwares embarcados.

Um teste realizado com a ferramenta STM32Cube.AI<sup>1</sup> permitiu estimar os requisitos de um microcontrolador para rodar os modelos testados. A Tabela 6 exhibe estes requisitos mínimos considerando uma plataforma ARM Cortex M0 ou superior, *AI compatible*. Os modelos leves propostos são todos capazes de serem executados em microcontroladores de 256 kB de RAM e memória Flash. O modelo proposto por Katsaouni et al. (2022) e a DWNet possuem os menores requisitos de memória e poderiam ser embarcados em

---

<sup>1</sup><https://www.st.com/en/embedded-software/x-cube-ai.html>

microcontroladores mais simples e econômicos. Por outro lado, os modelos ResNet-34 e VGG-11 necessitam de uma quantidade de memória Flash elevada, o que em muitos casos requisita um módulo de armazenamento externo ao microcontrolador. Experimentos com estes microcontroladores são necessários para estimar o tempo de inferência dos modelos compatíveis.

Tabela 6: Requisitos computacionais mínimos para execução dos modelos em microcontrolador ARM. \* indica modelo com frequência de amostragem de 200 Hz.

<b>Modelo</b>	<b>Flash</b>	<b>RAM</b>
<b>Katsaouni2022</b>	17,97 kB	37,96 kB
<b>DWNet</b>	38,34 kB	86,89 kB
<b>DWNet*</b>	38,34 kB	131,89 kB
<b>LiteVGG-11</b>	108,48 kB	151,00 kB
<b>LiteResNet-34</b>	221,71 kB	146,37 kB
<b>ResNet-34</b>	63,33 kB	291,67 kB
<b>VGG-11</b>	395,83kB	678,02 kB

Fonte: De autoria própria.

## 4.2 Classificação binária

Os resultados obtidos no experimento preliminar, em termos da média obtida na validação cruzada de Monte Carlo, estão detalhados na Tabela 13 do Apêndice C. Os modelos de Nurmaini et al. (2020) e LiteVGG-19 não convergiram.

Para o experimento mais robusto, foram selecionados os modelos com melhores F1-score e FNR, sendo que se uma mesma arquitetura figurar mais de uma vez entre os melhores, o modelo com maior custo computacional será desconsiderado. Foram selecionados os modelos com F1-score superior a 90% e FNR inferior a 10%: LiteVGG-11, LiteResNet-18, ResNet-18, VGG-11 e o modelo de Hsieh et al. (2020). O modelo DWNet, apesar de apresentar desempenho inferior, também foi selecionado para o experimento mais robusto, por ser um dos modelos aqui propostos. Treinamento e validação destes modelos foram realizados conforme método descrito na Seção 3.6.

As métricas obtidas são apresentadas na Tabela 7. Os resultados estão ordenados por ordem decrescente de F1-score. É apresentado o valor-P do teste t de Student comparando o F1-score do melhor modelo (em termos de F1-score) com os demais, sendo que os modelos sem diferença estatística para um  $\alpha = 5\%$  estão sombreados.



Tabela 7: Métricas obtidas para o experimento robusto de classificação binária.

Modelo	FS	AUC	ACC	PPV	TPR	TNR	FNR	F1-score	p-value
<b>LiteVGG-11</b>	200	0,9924	0,9691	0,9693	0,9689	0,9693	0,0311	0,9691	—
<b>LiteVGG-11</b>	128	0,9923	0,9691	0,9691	0,9690	0,9691	0,0310	0,9691	0,9614
<b>ResNet-18</b>	200	0,9916	0,9677	0,9635	0,9724	0,9631	0,0276	0,9679	0,1707
<b>LiteResNet-18</b>	200	0,9919	0,9677	0,9672	0,9683	0,9671	0,0317	0,9677	0,1638
<b>ResNet-18</b>	128	0,9917	0,9670	0,9635	0,9708	0,9632	0,0292	0,9671	0,0720
<b>LiteResNet-18</b>	128	0,9914	0,9654	0,9643	0,9667	0,9641	0,0333	0,9655	0,0005
<b>Hsieh</b>	200	0,9898	0,9636	0,9631	0,9643	0,9629	0,0357	0,9637	0,0001
<b>Hsieh</b>	128	0,9897	0,9628	0,9652	0,9604	0,9652	0,0396	0,9628	0,0001
<b>VGG-11</b>	128	0,9865	0,9595	0,9655	0,9532	0,9658	0,0468	0,9593	0,0000
<b>DWNet</b>	200	0,9850	0,9451	0,9483	0,9419	0,9483	0,0581	0,9450	0,0000
<b>VGG-11</b>	200	0,9577	0,9306	0,9406	0,9451	0,9161	0,0549	0,9385	0,0470
<b>DWNet</b>	128	0,9795	0,9345	0,9401	0,9282	0,9408	0,0718	0,9340	0,0000

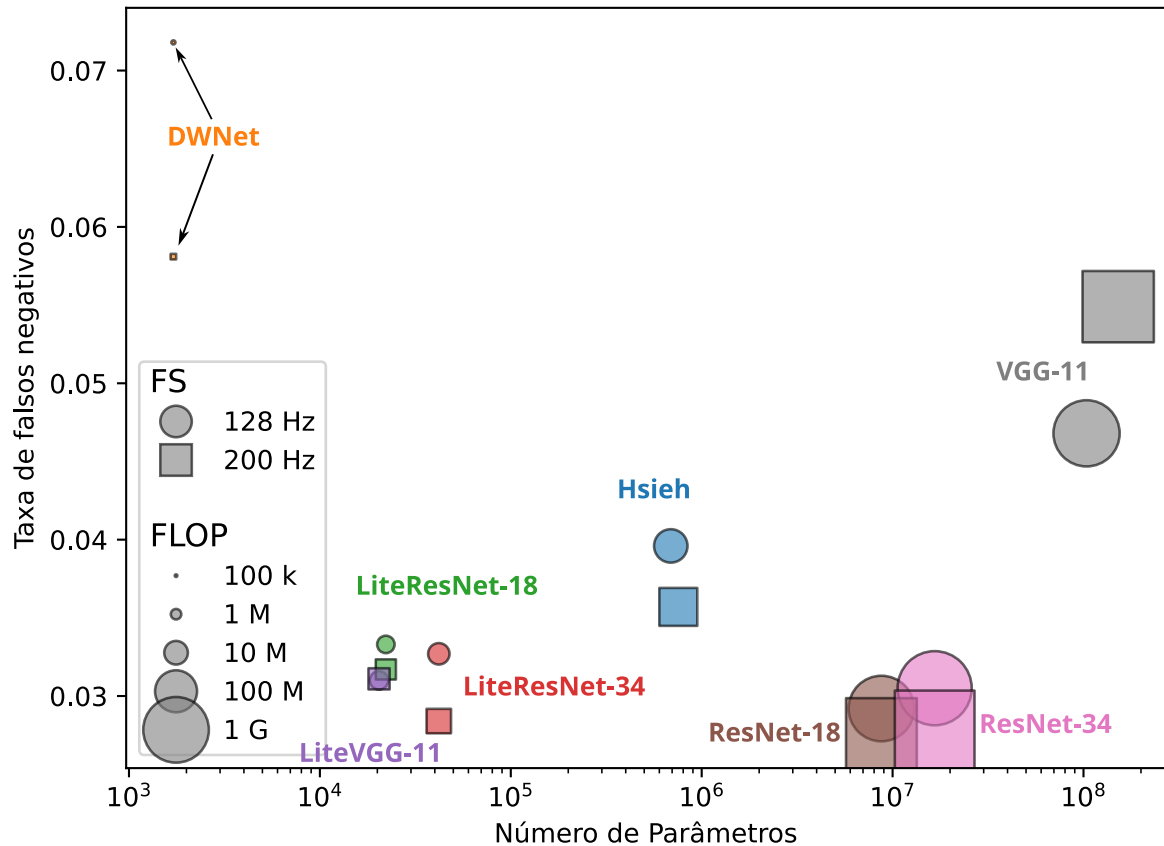
Fonte: De autoria própria.

O modelo LiteVGG-11, independente da frequência de amostragem, obteve o melhor F1-score médio ao longo das 20 amostragens da validação cruzada. Observa-se que para a classificação de um sinal de ECG em FA ou não-FA, considerando os pré-processamentos adotados, o modelo leve proposto foi capaz de se equiparar aos modelos mais custosos computacionalmente.

Como não foi notada uma diferença estatisticamente significativa entre os primeiros cinco modelos da Tabela 7, considerando o custo computacional, o modelo LiteVGG-11 com frequência de amostragem de 128 Hz pode ser considerado mais adequado para esta classificação em dispositivos móveis, já que requer menos custo computacional e opera com uma frequência de amostragem que pode resultar em redução do gasto energético do eletrocardiógrafo vestível.

A Figura 39 mostra a disposição dos modelos em relação a taxa de falsos positivos e custos computacionais (tanto em termos de número de parâmetros, eixo horizontal, quanto de FLOP, tamanho do marcador). Para esta métrica, os modelos ResNet-18, ResNet-34 e LiteResNet-34 apresentaram menores falsos negativos, entretanto à custa de maior complexidade computacional. Já os modelos LiteVGG-11 e LiteResNet-18 apresentam valores de falsos negativos próximos aos melhores modelos para esta métrica, mas com custo computacional reduzido. O modelo DWNet possui um custo computacional menor do que todos os demais modelos, mas este baixíssimo custo computacional é obtido à custa de um aumento de 3 a 4 pontos percentuais na taxa de falsos negativos.

Figura 39: Relação de custo computacional *versus* taxa de falsos negativos na classificação binária.



Fonte: De autoria própria.

### 4.3 Classificação multiclasse

Os resultados obtidos no experimento preliminar, em termos de métricas médias para a classe *FA*, estão detalhados na Tabela 14 do Apêndice C. Os modelos de Nurmaini et al. (2020) e LiteVGG-19 não convergiram.

Para o experimento mais robusto, do mesmo modo que no experimento de classificação binária, foram selecionados os modelos com melhores F1-score e FNR (para a classe *FA*), desconsiderando o modelo de maior custo computacional no caso de uma mesma arquitetura ser selecionada mais de uma vez entre os melhores. Foram selecionados os modelos com F1-score superior a 85% e FNR inferior a 10%, nas métricas associadas a classe *FA* em pelo menos uma configuração de frequência de amostragem. Novamente os modelos LiteVGG-11, LiteResNet-18, ResNet-18, VGG-11 e o modelo de Hsieh et al. (2020) atenderam os requisitos de corte foram os selecionados para o teste robusto. O modelo DWNNet novamente foi adicionado apesar do baixo desempenho. Treinamento e

validação destes modelos foram realizados conforme método descrito na Seção 3.6.

As métricas obtidas para a classe *FA* são apresentadas na Tabela 8, para a classe *RSN* na Tabela 9 e para a classe *outros* na Tabela 10. Os resultados em cada tabela estão ordenados por ordem decrescente de F1-score. O valor-P se refere ao teste t de Student comparando o F1-score do melhor modelo, para esta métrica, em cada uma das classes com os demais, sendo que os modelos sem diferença estatística ( $\alpha = 5\%$ ) estão sombreados.

Tabela 8: Métricas obtidas para a classe *FA* no experimento robusto de classificação multiclasse.

Modelo	FS	AUC	ACC	PPV	TPR	TNR	FNR	F1-score	p-value
LiteVGG-11	128	0,9911	0,8656	0,9458	0,9594	0,9725	0,0406	0,9525	—
ResNet-18	200	0,9908	0,8650	0,9540	0,9505	0,9770	0,0495	0,9521	0,8377
LiteVGG-11	200	0,9908	0,8624	0,9423	0,9597	0,9706	0,0403	0,9509	0,3604
LiteResNet-18	200	0,9905	0,8665	0,9507	0,9510	0,9753	0,0490	0,9508	0,3795
ResNet-34	128	0,9901	0,8601	0,9516	0,9482	0,9758	0,0518	0,9498	0,2437
ResNet-18	128	0,9898	0,8560	0,9520	0,9422	0,9762	0,0578	0,9470	0,0223
LiteResNet-34	128	0,9894	0,8580	0,9494	0,9440	0,9747	0,0560	0,9465	0,0121
LiteResNet-18	128	0,9895	0,8534	0,9465	0,9457	0,9732	0,0543	0,9460	0,0069
Hsieh	200	0,9872	0,8367	0,9335	0,9487	0,9661	0,0513	0,9409	0,0000
Hsieh	128	0,9865	0,8316	0,9320	0,9482	0,9653	0,0518	0,9399	0,0001
VGG-11	128	0,9837	0,8161	0,9217	0,9407	0,9599	0,0593	0,9310	0,0000
DWNet	200	0,9813	0,7910	0,8978	0,9146	0,9476	0,0854	0,9055	0,0000
VGG-11	200	0,8393	0,6662	0,6777	0,7830	0,8471	0,2170	0,7104	0,0149

Fonte: De autoria própria.

O modelo LiteVGG-11 apresentou o melhor desempenho na detecção de *FA* e *RSN*, entretanto, foi superado pelo modelo LiteResNet-18 na detecção da classe *Outros*, apresentando diferença significativa para um sinal de entrada amostrado a 200 Hz. Os resultados médios, apresentados na Tabela 11, indicam que nesta comparação também não houve diferença estatística entre os modelos LiteVGG-11 e LiteResNet-18. Como estes dois modelos, bem como o modelo ResNet-18 tradicional, apresentaram diferenças de F1-score inferior a um ponto percentual e não apresentaram diferença significativa na maioria das comparações, não é possível definir qual dos três é melhor neste quesito.

Por outro lado, como o foco é a classificação da *FA*, o modelo LiteVGG-11 apresentou as menores taxas de falsos negativos (em ambas configurações de frequência de amostragem), sendo esta diferença significativa em relação a todos os demais modelos (p-valor  $< 0,05$ ). O segundo melhor modelo neste quesito, o LiteResNet-18, apresenta cerca de 0,85 pontos percentuais a mais de falso diagnóstico.

Tabela 9: Métricas obtidas para a classe *RSN* no experimento robusto de classificação multiclasse.

Modelo	FS	AUC	ACC	PPV	TPR	TNR	FNR	F1-score	p-value
LiteVGG-11	128	0,9574	0,8656	0,8175	0,8712	0,9024	0,1288	0,8431	—
LiteResNet-18	200	0,9566	0,8665	0,8338	0,8528	0,9146	0,1472	0,8426	0,8668
LiteVGG-11	200	0,9555	0,8624	0,8126	0,8699	0,8994	0,1301	0,8401	0,2363
ResNet-18	200	0,9557	0,8650	0,8312	0,8470	0,9135	0,1530	0,8383	0,2033
LiteResNet-34	128	0,9531	0,8580	0,8303	0,8345	0,9141	0,1655	0,8318	0,0000
ResNet-18	128	0,9528	0,8560	0,8194	0,8464	0,9050	0,1536	0,8310	0,0149
ResNet-34	128	0,9538	0,8601	0,8404	0,8223	0,9214	0,1777	0,8305	0,0104
LiteResNet-18	128	0,9506	0,8534	0,8246	0,8256	0,9115	0,1744	0,8240	0,0106
Hsieh	200	0,9405	0,8367	0,7959	0,8213	0,8934	0,1787	0,8073	0,0000
Hsieh	128	0,9371	0,8316	0,7958	0,8040	0,8952	0,1960	0,7982	0,0000
VGG-11	128	0,9282	0,8161	0,7864	0,7760	0,8938	0,2240	0,7804	0,0000
DWNet	200	0,9168	0,7910	0,7592	0,7734	0,8745	0,2266	0,7646	0,0000
VGG-11	200	0,7958	0,6662	0,5863	0,6576	0,8033	0,3424	0,6014	0,0044

Fonte: De autoria própria.

Tabela 10: Métricas obtidas para a classe *Outros* no experimento robusto de classificação multiclasse.

Modelo	FS	AUC	ACC	PPV	TPR	TNR	FNR	F1-score	p-value
LiteResNet-18	200	0,9352	0,8665	0,8165	0,7958	0,9100	0,2042	0,8055	—
ResNet-18	200	0,9349	0,8650	0,8121	0,7975	0,9069	0,2025	0,8041	0,6731
ResNet-34	128	0,9320	0,8601	0,7925	0,8097	0,8929	0,1903	0,8001	0,1618
LiteVGG-11	128	0,9366	0,8656	0,8344	0,7662	0,9235	0,2338	0,7984	0,0302
LiteResNet-34	128	0,9302	0,8580	0,7968	0,7954	0,8982	0,2046	0,7956	0,0070
LiteVGG-11	200	0,9344	0,8624	0,8324	0,7575	0,9235	0,2425	0,7929	0,0010
LiteResNet-18	128	0,9289	0,8534	0,7928	0,7890	0,8954	0,2110	0,7898	0,0012
ResNet-18	128	0,9289	0,8560	0,8039	0,7795	0,9029	0,2205	0,7892	0,0290
Hsieh	200	0,9120	0,8367	0,7815	0,7401	0,8955	0,2599	0,7589	0,0000
Hsieh	128	0,9097	0,8316	0,7693	0,7427	0,8869	0,2573	0,7537	0,0000
VGG-11	128	0,8910	0,8161	0,7398	0,7315	0,8704	0,2685	0,7346	0,0000
DWNet	200	0,8724	0,7910	0,7163	0,6850	0,8644	0,3150	0,6991	0,0000
VGG-11	200	0,7660	0,6662	0,5362	0,5580	0,8489	0,4420	0,5364	0,0026

Fonte: De autoria própria.

No geral, os modelos de menor custo computacional das arquiteturas LiteVGG e LiteResNet apresentaram as melhores métricas, superando as suas contrapartes originais e também os modelos apresentados na literatura.

Tabela 11: Métricas média entre todas as classes no experimento robusto de classificação multiclasse.

Modelo	FS	AUC	ACC	PPV	TPR	TNR	FNR	F1-score	p-value
LiteResNet-18	200	0,9608	0,8665	0,8670	0,8665	0,9333	0,1335	0,8663	—
ResNet-18	200	0,9605	0,8650	0,8657	0,8650	0,9325	0,1350	0,8648	0,5721
LiteVGG-11	128	0,9617	0,8656	0,8659	0,8656	0,9328	0,1344	0,8647	0,4955
LiteVGG-11	200	0,9602	0,8624	0,8624	0,8624	0,9312	0,1376	0,8613	0,0413

Fonte: De autoria própria.

## 4.4 Classificação multilabel

Os resultados obtidos no experimento preliminar, em termos de métricas médias para a classe *FA*, estão detalhados na Tabela 15 do Apêndice C. O modelo de Nurmaini et al. (2020) não convergiu em nenhuma configuração de frequência de amostragem.

Para o experimento mais robusto, do mesmo modo que nos experimentos anteriores, foram selecionados os modelos com melhores F1-score e FNR (para a classe *FA*), entretanto, desta vez não foi desconsiderado o modelo de maior custo computacional no caso de uma mesma arquitetura ser selecionada mais de uma vez entre os melhores, isto porque o modelo ResNet-34 apresentou F1-score superior ao ResNet-18 em quase um ponto percentual e taxa de falsos negativos mais de dois pontos inferior. Foram selecionados os modelos com F1-score superior a 85% e FNR inferior a 10%, nas métricas associadas a classe *FA* em pelo menos uma configuração de frequência de amostragem. Os modelos ResNet-34, LiteResNet-34, LiteVGG-11, LiteResNet-18, ResNet-18 e o modelo de Hsieh et al. (2020) atenderam os requisitos de corte e foram os selecionados para o teste robusto. O modelo VGG-11 também foi incluído, mesmo apresentando desempenho inferior ao corte estabelecido, para possibilitar comparação direta com o LiteVGG-11. Treinamento e validação destes modelos foram realizados conforme método descrito na Seção 3.6.

Diante do volume de dados resultantes deste experimento, os resultados completos se encontram na Tabela 16 apresentada no Apêndice D. A Tabela 12 apresenta um recorte dos resultados, focando nos modelos ResNet-34, ResNet-18, LiteResNet-34 e LiteResNet-18 utilizando a frequência de amostragem de 200 Hz, e nos modelos LiteVGG-11 e VGG-11 para a frequência de amostragem de 128 Hz. Estes modelos com estas configurações de frequência de amostragem foram escolhidos por serem as configurações que obtiveram melhores F1-score, com exceção do VGG-11 que foi incluído para comparação direta com o LiteVGG-11.

Tabela 12: Melhores métricas obtidas no experimento robusto de classificação multilabel. Melhor resultados para cada métrica/classe está destacado em negrito. Métricas sombreadas indicam falta diferença estatística ( $\alpha = 5\%$ ) em relação à melhor métrica.

Modelo	FS	Classe	Métricas por classe			Métricas médias			
			TPR	TNR	F1-score	AUC	ACC	TPR	TNR
LiteVGG-11	128	FA	<b>0.9442</b>	0.9839	0.9408	0.9890	0.8740	<b>0.9346</b>	0.9733
		RSN	<b>0.8717</b>	0.9485	0.8406				
		BS	0.9623	0.9785	0.9402				
		TS	0.9604	0.9824	0.9464				
VGG-11	128	FA	0.9143	0.9749	0.9084	0.9813	0.8337	0.8863	0.9715
		RSN	0.7337	0.9514	0.7619				
		BS	0.9559	0.9771	0.9344				
		TS	0.9414	0.9825	0.9366				
LiteResNet-18	200	FA	0.9357	0.9856	0.9394	0.9887	0.8755	0.9265	0.9762
		RSN	0.8546	0.9542	0.8400				
		BS	0.9570	0.9818	0.9435				
		TS	0.9589	0.9832	0.9470				
ResNet-18	200	FA	0.9251	<b>0.9896</b>	<b>0.9411</b>	0.9886	<b>0.8783</b>	0.9213	<b>0.9784</b>
		RSN	0.8503	<b>0.9580</b>	<b>0.8439</b>				
		BS	0.9508	<b>0.9835</b>	0.9435				
		TS	0.9589	<b>0.9826</b>	0.9460				
LiteResNet-34	200	FA	0.9375	0.9856	0.9404	0.9889	0.8791	0.9279	0.9770
		RSN	0.8515	0.9575	0.8437				
		BS	<b>0.9629</b>	0.9817	<b>0.9465</b>				
		TS	0.9598	0.9832	<b>0.9475</b>				
ResNet-34	200	FA	0.9266	<b>0.9890</b>	0.9408	0.9889	0.8770	0.9218	0.9779
		RSN	0.8469	0.9569	0.8400				
		BS	0.9531	0.9830	0.9438				
		TS	<b>0.9606</b>	0.9828	0.9472				

Fonte: De autoria própria.

Nota-se que o modelo LiteVGG-11 apresentou as melhores taxas de verdadeiro positivo (sensibilidade) ao considerar o conjunto de todas as classes. Isto se reflete em menor taxa de falsos negativos. A sensibilidade do modelo LiteVGG-11 para as classes *FA* e *RSN* apresentou diferença estatística significativa em relação a todos os demais modelos, concretizando a sua superioridade neste quesito.

Considerando a métrica de taxa de falsos negativos (especificidade), os modelos mais complexos computacionalmente, ResNet-18 e ResNet-34, apresentaram clara superioridade em relação aos demais modelos.

Os modelos LiteResNet-34 e ResNet-34 apresentaram melhores desempenhos em relação ao F1-score em todas as classes, seguido pelos modelos LiteVGG-11, LiteResNet-18 e

ResNet-18. Estes últimos apresentaram um degradação em relação aos primeiros apenas na classe de bradicardia sinusal.

Vale ressaltar que, apesar de existirem diferenças estatísticas entre os modelos LiteVGG-11, LiteResNet-18, LiteResNet-34, ResNet-18 e ResNet-34 para algumas métricas em determinadas classes, a diferença numérica é muito baixa e os modelos propostos apresentaram desempenho similar ou superior às suas contrapartes originais. Entretanto, novamente considerando a situação de monitoramento contínuo do ritmo cardíaco por meio de eletrocardiógrafos vestíveis, em especial para pacientes com risco de desenvolvimento eventos de fibrilação, o modelo LiteVGG-11 apresenta superioridade devido ao seu baixo custo computacional e menor taxa de falsos negativos para FA mesmo em sinais adquiridos com frequência de amostragem de apenas 128 Hz.

## 4.5 Discussões da Classificação

Levando em consideração o objetivo principal deste trabalho, que é construir um modelo de baixo custo computacional para a classificação de FA, o modelo LiteVGG-11 apresentou certa superioridade nesta tarefa enquanto manteve um dos menores custos computacionais dentre os modelos aqui estudados.

Em relação às estratégias de classificação, a classificação binária resultou em melhor discriminação da FA, apresentando F1-score 1,66 e 2,83 pontos percentuais superior às classificações multiclasse e multilabel, respectivamente, e taxas de falsos negativos 0,96 e 2,48 pontos percentuais inferiores a tais estratégias. Tal resultado já era esperado e advém do fato de que durante o treinamento para a classificação binária, todos os pesos do modelo são otimizados apenas para a detecção desta classe, enquanto nas demais estratégias as camadas convolucionais possuem pesos que são otimizados simultaneamente para todas as classes treinadas, sendo que apenas a última camada densa de saída apresenta uma otimização específica para cada classe.

Dada a superioridade da classificação binária para a identificação da FA, os sinais identificados como não-FA podem ser posteriormente estratificados em taquicardia, bradicardia e frequência cardíaca normal por meio de algoritmos clássicos de estimativa de frequência cardíaca, como o algoritmo de Pan-Tompkins (PAN; TOMPKINS, 1985). Entretanto esta estratificação também possui um certo custo computacional, sendo assim, o uso do classificador multiclasse ou multilabel pode ser justificado em algumas situações de baixa capacidade computacional e baixos requisitos de acurácia no diagnóstico.

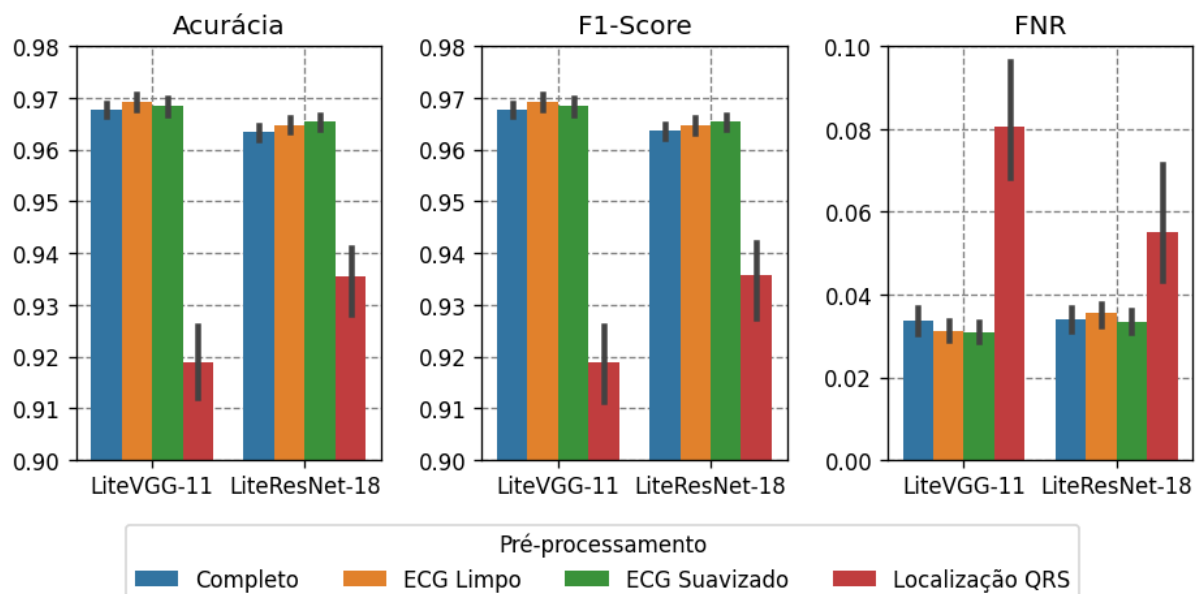
## 4.6 Impacto do pré-processamento

Aproveitando o experimento robusto de classificação multiclasse, também foi realizado um estudo do impacto do pré-processamento proposto. Neste estudo foram treinados modelos com entrada apenas do sinal de ECG limpo (1-40 Hz), do sinal de ECG suavizado (1-15 Hz) e do sinal de localização dos complexos QRS. Por simplicidade, este experimento foi realizado considerando apenas os modelos LiteVGG-11 e LiteResNet-34, ambos com frequência de amostragem de 128 Hz. A figura 40 apresenta os resultados obtidos em uma validação cruzada de Monte-Carlo com 20 amostragens.

O pré-processamento proposto apresentou desempenho inferior à utilização do ECG limpo ou do ECG suavizado. Entretanto, este pré-processamento foi desenvolvido considerando um caso de elevado nível de ruído, como aquele encontrado em monitoramentos contínuos de ritmo cardíaco, onde o paciente utiliza um ECG vestível durante suas atividades cotidianas. Os datasets utilizados neste trabalho, no entanto, foram adquiridos em ambiente clínico, onde diversas fontes de ruídos são minimizadas.

Para simular de forma simplificada um ambiente contendo elevado nível de ruído, os modelos, já treinados com os pré-processamentos analisados, foram testados em sinais com adição controlada de ruído branco. Ou seja, os sinais de testes de cada uma das amostragens da validação cruzada foram contaminados em determinados valores de SNR antes

Figura 40: Impacto do pré-processamento nas métricas da classificação binária. Barras representam valor médio e traços indicam o intervalo de confiança a 95%.



Fonte: De autoria própria.

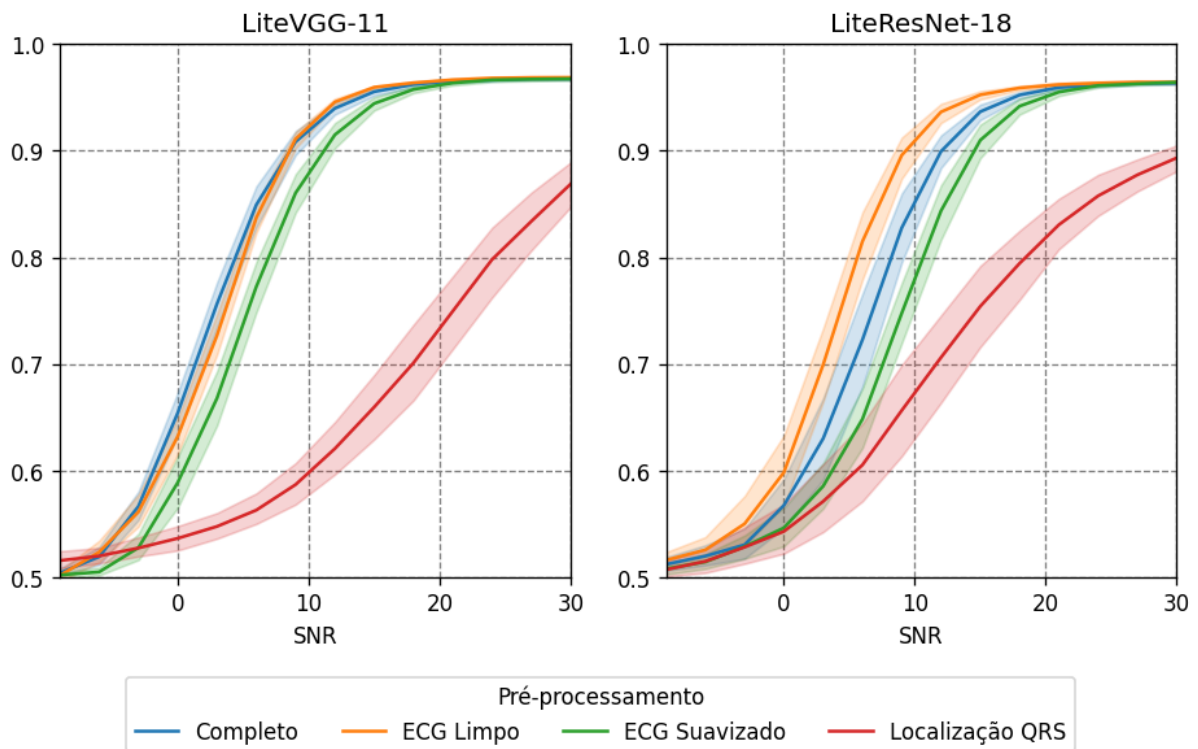


de serem pré-processados e classificados para teste. A Figura 41 apresenta os resultados obtidos nesta análise.

Para o modelo LiteResNet-18, o pré-processamento adotado interferiu negativamente em sua acurácia, independentemente do nível de ruído do sinal. Já para o modelo LiteVGG-11, ambos pré-processamento completo e ECG limpo atingiram acurácias similares ao longo da faixa de SNR analisada, sendo que o pré-processamento completo apresentou acurácia média superior para SNR inferior a 8 dB.

Uma aplicação clínica do monitoramento de ritmo cardíaco deverá levar em consideração o nível de ruído. Caso o hardware de aquisição possua boas estratégias de mitigação de ruído e artefatos, fornecendo um sinal de alta qualidade, o pré-processamento tradicional (ECG Limpo) poderia ser utilizado, reduzindo ainda mais o número de FLOP para inferência, e conseqüentemente os custos computacionais. Também vale ressaltar que a análise aqui proposta é simples e não leva em consideração os tipos de ruídos específicos do ECG, apenas ruído branco.

Figura 41: Acurácia da classificação binária de acordo com a SNR simulada. Região sombreada indica o intervalo de confiança a 95%.



Fonte: De autoria própria.

## 5 CONCLUSÃO

Os experimentos conduzidos neste trabalho indicam que os modelos de aprendizado profundo existentes já são capazes de serem executados em plataforma móvel, com tempo de inferência baixo o suficiente para realizar pelo menos uma medida por segundo. Os modelos de baixo custo computacional propostos apresentaram tempo de inferência cerca de uma ordem de grandeza inferior aos modelos existentes, além de terem cerca de 2 a 4 ordens de grandeza menos parâmetros e custo em termos de operações de ponto flutuante. Os modelos propostos possuem custo computacional baixo o suficiente para serem inclusive executados em microcontroladores ARM Cortex M0 compatíveis com IA.

A estratégia adotada de utilização de convoluções separáveis, menor número de filtros por convolução, *poolings* globais e menos unidades nas camadas densas se mostrou efetiva não somente na redução do custo computacional total dos modelos, como também na manutenção de seus desempenhos de classificação, conforme pode ser apurado nos experimentos de classificação binária, multiclasse e multilabel propostos. Esta estratégia de adaptação proposta inclusive deu origem à arquitetura LiteVGG, que se mostrou superior à arquitetura VGGNet na classificação de ritmos cardíacos através de sinais de ECG. O modelo mais leve proposto, a DWNet, possui custo computacional inferior aos demais modelos, mas não atingiu métricas similares aos demais modelos, apesar de superar o modelo mais leve encontrado na literatura.

O pré-processamento proposto apresentou vantagens apenas em situações de SNR inferior a 8 dB e apenas para o modelo LiteVGG-11, resultando em perdas de acurácia em relação a um pré-processamento mais tradicional em outras condições. Entretanto, visto que a simulação de ruído aqui utilizada consistia de uma simples adição de ruído branco, ainda se faz necessário verificar o desempenho de tal pré-processamento em sinais com ruídos e artefatos reais, como aqueles coletados com eletrocardiógrafos vestíveis em condições de uso diário.

## 5.1 Contribuições

Neste trabalho foram propostos e avaliados modelo de aprendizado profundo para a classificação de ritmos cardíacos considerando a possibilidade de seu uso em sinais adquiridos com eletrocardiógrafos vestíveis. Com estes modelos se torna possível desenvolver uma plataforma de monitoramento de ritmo cardíaco para pacientes de alto risco para eventos de fibrilação.

A vantagem dos modelos desenvolvidos, em relação aos já presentes na literatura, está no baixo requisito computacional, o que possibilitaria a execução da plataforma em smartphones de entrada, ampliando assim a população que poderia fazer uso desta plataforma.

## REFERÊNCIAS

- ABADI, M. et al. Tensorflow: A system for large-scale machine learning. In: Proceedings of the 12th USENIX Conference on Operating Systems Design and Implementation, 12th., Savannah, GA, USA. **OSDI 16**. [S.l.]: USENIX Association, p. 265–283. ISBN 9781931971331, 2016. DOI: 10.48550/arXiv.1605.0869.
- ADH30 et al. **A Wiggers diagram, showing the cardiac cycle events occurring in the left ventricle**. Wikimedia Commons, the free media repository, 2016. Disponível em: [https://commons.wikimedia.org/wiki/File:Wiggers\\_Diagram\\_2.svg](https://commons.wikimedia.org/wiki/File:Wiggers_Diagram_2.svg). Acesso em: 22 oct. 2023.
- AHMED, N.; ZHU, Y. Early detection of atrial fibrillation based on ECG signals. **Bioengineering**, v. 7, n. 1, p. 16, fev. 2020. DOI: 10.3390/bioengineering7010016.
- AMIN, A. S.; TAN, H. L.; WILDE, A. A. Cardiac ion channels in health and disease. **Heart Rhythm**, v. 7, n. 1, p. 117–126, 2010. DOI: 10.1016/j.hrthm.2009.08.005.
- ANDERSON, R. H. et al. The anatomy of the cardiac conduction system. **Clinical Anatomy**, v. 22, n. 1, p. 99–113, 2009. DOI: 10.1002/ca.20700.
- ARANKI, S. F. et al. Predictors of atrial fibrillation after coronary artery surgery. **Circulation**, v. 94, n. 3, p. 390–397, ago. 1996. DOI: 10.1161/01.cir.94.3.390.
- ATKIELSKI, A. **ECG of a heart in normal sinus rhythm**. Wikimedia Commons, the free media repository, 2022. Disponível em: <https://commons.wikimedia.org/wiki/File:SinusRhythmLabels.svg>. Acesso em: 18 oct. 2023.
- BABAR, T. B. Atrial fibrillation. In: \_\_\_\_\_. **Ferri’s Clinical Advisor 2022**. 1. ed. Philadelphia, USA: Elsevier, 2022. p. 204–214. ISBN 978-0-323-75889-5.
- BANGKOK DUSIT MEDICAL SERVICES. **Electrocardiogram (EKG)**. 2022. Disponível em: <https://www.bangkokhearthospital.com/en/medical-service/electrocardiogram-ekg>. Acesso em: 14 jun. 2023.
- BOUSSELJOT, R.; KREISELER, D.; SCHNABEL, A. Nutzung der EKG-signaldatenbank CARDIODAT der PTB über das internet. **Biomedizinische Technik/Biomedical Engineering**, v. 40, n. s1, p. 317–318, jul 1995. DOI: 10.1515/bmte.1995.40.s1.317.
- BOUZID, Z. et al. Remote and wearable ECG devices with diagnostic abilities in adults: A state-of-the-science scoping review. **Heart Rhythm**, v. 19, n. 7, p. 1192–1201, 2022. DOI: 10.1016/j.hrthm.2022.02.030.

- BROWN, T. B. et al. Language models are few-shot learners. In: Proceedings of the 34th International Conference on Neural Information Processing Systems, Vancouver, BC, Canada. **NIPS 2020**. Red Hook, NY, USA: Curran Associates Inc. ISBN 9781713829546, 2020. DOI: 10.48550/ARXIV.2005.14165.
- CADOGAN, M. **ECG lead positioning**. 2022. Disponível em: <https://litfl.com/ecg-lead-positioning/>. Acesso em: 19 oct. 2023.
- CAMANHO, L. E. M.; SANTOS, G. V. d. Postoperative atrial fibrillation: The challenge of risk prediction. **International Journal of Cardiovascular Sciences**, v. 32, n. 6, p. 594–595, Nov 2019. DOI: 10.5935/2359-4802.20190088.
- CARMO, R. L. **Esqueleto cardíaco**. Kenhub, 2022. Disponível em: <https://www.kenhub.com/pt/library/anatomia/esqueleto-cardiaco>. Acesso em: 19 oct. 2023.
- CHEN, J. Atrial fibrillation and atrial flutter. **Clinics in Geriatric Medicine**, v. 28, n. 4, p. 635–647, nov 2012. DOI: 10.1016/j.cger.2012.08.002.
- CHOLLET, F. Xception: Deep learning with depthwise separable convolutions. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition, 21-26 Jul 2016, Honolulu, HI, USA. **CVPR 2017**. [S.l.]: IEEE Computer Society, p. 1800–1807. ISSN 1063-6919, jul 2017. DOI: 10.1109/CVPR.2017.195.
- CHOLLET, F. et al. **Keras**. 2015. Disponível em: <https://keras.io>. Acesso em: 09 fev. 2022.
- CMGLEE; THOMA, M. **Receiver Operating Characteristic (ROC) curve with False Positive Rate and True Positive Rate**. Wikimedia Commons, the free media repository, 2021. Disponível em: [https://commons.wikimedia.org/wiki/File:Roc\\_curve.svg](https://commons.wikimedia.org/wiki/File:Roc_curve.svg). Acesso em: 26 oct. 2023.
- COMPANY-BOSCH, E.; HARTMANN, E. **ECG front-end design is simplified with MicroConverter®**. Analog Devices, 2003. Disponível em: <https://www.analog.com/en/analog-dialogue/articles/ecg-front-end-design-simplified.html>. Acesso em: 19 oct. 2023.
- COSÍO, F. G. et al. Atrial flutter: an update. **Revista Española de Cardiología (English Edition)**, v. 59, n. 8, p. 816–831, jan 2006. DOI: 10.1016/s1885-5857(07)60047-4.
- CRANDALL, M. A. et al. Atrial fibrillation significantly increases total mortality and stroke risk beyond that conveyed by the CHADS2 risk factors. **Pacing and clinical electrophysiology**, v. 32, n. 8, p. 981–986, 2009. DOI: 10.1111/j.1540-8159.2009.02427.x.
- DAVE, S. et al. Incidence and risk factors for development of atrial fibrillation after cardiac surgery under cardiopulmonary bypass. **Indian Journal of Anaesthesia**, v. 62, n. 11, p. 887, 2018. DOI: 10.4103/ija.ija618.
- DIAS, F. M.; RIBEIRO, E.; SOARES, Q.; KRIEGER, J.; GUTIERREZ, M. A. Predicting age and sex from reduced lead electrocardiograms using deep learning. **TechRxiv**, Outubro 2023. DOI: 10.36227/techrxiv.23908089.v1.
- DIFRANCESCO, D.; BORER, J. S. The funny current. **Drugs**, v. 67, n. Supplement 2, p. 15–24, 2007. DOI: 10.2165/00003495-200767002-00003.

- FUJITA, H.; CIMR, D. Computer aided detection for fibrillations and flutters using deep convolutional neural network. **Information Sciences**, v. 486, p. 231–239, 2019. DOI: 10.1016/j.ins.2019.02.065.
- FURST, B. The heart: Pressure-propulsion pump or organ of impedance? **Journal of Cardiothoracic and Vascular Anesthesia**, v. 29, n. 6, p. 1688–1701, Dec 2015. DOI: 10.1053/j.jvca.2015.02.022.
- GALLI, A.; AMBROSINI, F.; LOMBARDI, F. Holter monitoring and loop recorders: From research to clinical practice. **Arrhythmia & electrophysiology review**, v. 5, n. 2, p. 136, 2016. DOI: 10.15420/aer.2016.17.2.
- GARGIULO, G. D. True unipolar ECG machine for Wilson central terminal measurements. **BioMed Research International**, v. 2015, p. 586397, Oct 2015. DOI: 10.1155/2015/586397.
- GLOBAL BURDEN OF DISEASE COLLABORATIVE NETWORK. **Global Burden of Disease Study 2019 (GBD 2019)**. Seattle, United States: Institute for Health Metrics and Evaluation (IHME), 2020. Disponível em: <https://vizhub.healthdata.org/gbd-results/>.
- GOLDBERGER, A. L.; GOLDBERGER, Z. D.; SHVILKIN, A. Chapter 13 - sinus and escape rhythms. In: GOLDBERGER, A. L.; GOLDBERGER, Z. D.; SHVILKIN, A. (Ed.). **Goldberger's Clinical Electrocardiography (Ninth Edition)**. 9. ed. [S.l.]: Elsevier, 2018. p. 122–129. ISBN 978-0-323-40169-2. DOI: 10.1016/B978-0-323-40169-2.00013-5.
- GOLDBERGER, A. L. et al. Physiobank, physiotoolkit, and physionet. **Circulation**, v. 101, n. 23, p. e215–e220, 2000. DOI: 10.1161/01.CIR.101.23.e215.
- GUYTON, A. C.; HALL, J. **Tratado de Fisiologia Médica**. 12. ed. Rio de Janeiro: Elsevier, 2011.
- HAFEEZ, Y.; GROSSMAN, S. A. **Sinus Bradycardia**. Treasure Island (FL): StatPearls Publishing, 2022. Disponível em: <http://europepmc.org/books/NBK493201>. Acesso em: 14 oct. 2023.
- HE, K. et al. Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition, 27-30 Jun 2016, Las Vegas, NV, USA. **CVPR 2016**. [S.l.], p. 770–778, 2016. DOI: 10.1109/CVPR.2016.90.
- HOWARD, A. G.; ZHU, M.; CHEN, B. et al. **MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications**. [S.l.]: arXiv, 2017. DOI: 10.48550/ARXIV.1704.04861.
- HSIEH, C.-H. et al. Detection of atrial fibrillation using 1d convolutional neural network. **Sensors**, v. 20, n. 7, p. 2136, apr 2020. DOI: 10.3390/s20072136.
- IKONNIKOV, G.; WONG, E. **Physiology of cardiac conduction and contractility - Action potential of cardiac muscles**. 2013. Disponível em: <http://www.pathophys.org/physiology-of-cardiac-conduction-and-contraction/actionpotential/>.

- IOFFE, S.; SZEGEDY, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In: BACH, F.; BLEI, D. (Ed.). **Proceedings of the 32nd International Conference on Machine Learning**. Lille, France: PMLR, v. 37, p. 448–456, 07–09 Jul 2015. DOI: 10.48550/arXiv.1502.03167.
- ISSA, Z. F.; MILLER, J. M.; ZIPES, D. P. Atrial fibrillation. In: ISSA, Z. F.; MILLER, J. M.; ZIPES, D. P. (Ed.). **Clinical Arrhythmology and Electrophysiology**. Philadelphia: W.B. Saunders, 2009. cap. 11, p. 208–286. DOI: 10.1016/b978-1-4160-5998-1.00014-8.
- JOUNG, S. et al. Early detection of atrial fibrillation-atrial flutter using remote patient monitoring. **Journal of Medical Cases**, v. 10, n. 2, p. 31–36, 2019. DOI: 10.14740/jmc3151w.
- KATSAOUNI, N. et al. Energy efficient convolutional neural networks for arrhythmia detection. **Array**, v. 13, p. 100127, mar. 2022. DOI: 10.1016/j.array.2022.100127.
- KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. In: International Conference on Learning Representations, 3rd., 7–9 May 2015, San Diego. **ICLR 2015**. [S.l.], 2015. Conference Track Proceedings.
- KIRCHHOF, P. et al. Early rhythm-control therapy in patients with atrial fibrillation. **New England Journal of Medicine**, v. 383, n. 14, p. 1305–1316, oct 2020. DOI: 10.1056/nejmoa2019422.
- KLIGFIELD, P. et al. Recommendations for the standardization and interpretation of the electrocardiogram. **Heart Rhythm**, v. 4, n. 3, p. 394–412, Mar 2007. DOI: 10.1016/j.hrthm.2007.01.027.
- KORNEJ, J. et al. Epidemiology of atrial fibrillation in the 21st century: Novel methods and new insights. **Circulation research**, v. 127, n. 1, p. 4–20, jun 2020. DOI: 10.1161/CIRCRESAHA.120.316340.
- LIN, M.; CHEN, Q.; YAN, S. Network in network. In: International Conference on Learning Representations, 2nd., 14–16 Apr 2014, Banff, AB, Canada. **ICLR 2014**. [S.l.], 2014. Conference Track Proceedings. DOI: 10.48550/arXiv.1312.4400.
- LIU, F. et al. An open access database for evaluating the algorithms of electrocardiogram rhythm and morphology abnormality detection. **Journal of Medical Imaging and Health Informatics**, v. 8, n. 7, p. 1368–1373, sep 2018. DOI: 10.1166/jmihi.2018.2442.
- LOPES, S. G. B. **Bio: volume único**. São Paulo, Brasil: Saraiva, 2004.
- Madhero88. **Electrical conduction system of the heart**. Wikimedia Commons, the free media repository, 2010. Disponível em: <https://commons.wikimedia.org/wiki/File:Conductionsystemoftheheart.svg>. Acesso em: 18 oct. 2023.
- MAGALHÃES, L. et al. II diretrizes brasileiras de fibrilação atrial. **Arquivos Brasileiros de Cardiologia**, v. 106, n. 4, 2016. DOI: 10.5935/abc.20160055.
- MAGNI, S. et al. Combining resnet model with handcrafted temporal features for ECG classification with varying number of leads. In: Computing in Cardiology, 48th., 12–15 Sep 2021, Brno, Czech Republic. **CINCC 2021**. [S.l.], v. 48, p. 1–4, 2021. DOI: 10.23919/CinC53138.2021.9662931.

- MAR, B. D. The history of clinical holter monitoring. **Annals of Noninvasive Electrocardiology**, v. 10, n. 2, p. 226–230, abr. 2005. DOI: 10.1111/j.1542-474x.2005.10202.x.
- MARROUCHE, N. F. et al. Catheter ablation for atrial fibrillation with heart failure. **New England Journal of Medicine**, v. 378, n. 5, p. 417–427, feb 2018. DOI: 10.1056/nejmoa1707855.
- MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **The bulletin of mathematical biophysics**, v. 5, n. 4, p. 115–133, Dec 1943. DOI: 10.1007/BF02478259.
- MEZOO. **HiCardi® System**. Mezo Co., Ltd., 2022. Disponível em: <https://www.me-zoo.com/main/html.php?htmid=products/hicardisystem.html>. Acesso em: 14 jun. 2023.
- MOEINZADEH, H. et al. WCTECGdb: A 12-lead electrocardiography dataset recorded simultaneously with raw exploring electrodes' potential directly referred to the right leg. **Sensors**, v. 20, n. 11, 2020. DOI: 10.3390/s20113275.
- MOODY, G. B.; MULDROW, W.; MARK, R. G. A noise stress test for arrhythmia detectors. **Computers in cardiology**, v. 11, n. 3, p. 381–384, 1984.
- NATTEL, S. Atrial electrophysiology and mechanisms of atrial fibrillation. **Journal of Cardiovascular Pharmacology and Therapeutics**, v. 8, n. 1\_suppl, p. S5–S11, 2003. DOI: 10.1177/107424840300800102.
- NEJEDLY, P. et al. Classification of ECG using ensemble of residual CNNs with attention mechanism. In: Computing in Cardiology, 48th., 12-15 Sep 2021, Brno, Czech Republic. **CINC 2021**. [S.l.], v. 48, p. 1–4, 2021. DOI: 10.23919/CinC53138.2021.9662723.
- NERBONNE, J. M.; KASS, R. S. Molecular physiology of cardiac repolarization. **Physiological Reviews**, v. 85, n. 4, p. 1205–1253, 2005. DOI: 10.1152/physrev.00002.2005.
- NOGUÉ, J. M. i. **Diagram of the human heart**. Wikimedia Commons, the free media repository, 2022. Disponível em: [https://commons.wikimedia.org/w/index.php?title=File:Diagram\\_of\\_the\\_human\\_heart\\_\(cropped\)\\_pt.svg](https://commons.wikimedia.org/w/index.php?title=File:Diagram_of_the_human_heart_(cropped)_pt.svg). Acesso em: 25 jul. 2022.
- NPATCHETT. **Limb leads of EKG**. Wikimedia Commons, the free media repository, 2021. Disponível em: [https://commons.wikimedia.org/w/index.php?title=File:Limb\\_leads\\_of\\_EKG.png](https://commons.wikimedia.org/w/index.php?title=File:Limb_leads_of_EKG.png). Acesso em: 19 oct. 2023.
- NURMAINI, S. et al. Robust detection of atrial fibrillation from short-term electrocardiogram using convolutional neural networks. **Future Generation Computer Systems**, v. 113, p. 304–317, 2020. DOI: 10.1016/j.future.2020.07.021.
- PAN, J.; TOMPKINS, W. J. A real-time QRS detection algorithm. **IEEE Transactions on Biomedical Engineering**, BME-32, n. 3, p. 230–236, 1985. DOI: 10.1109/TBME.1985.325532.
- PASSANNANTE, A. N. Prevention of atrial fibrillation after cardiac surgery. **Current Opinion in Anaesthesiology**, v. 24, n. 1, p. 58–63, fev. 2011. DOI: 10.1097/aco.0b013e3283414243.



PASZKE, A.; et al. Pytorch: An imperative style, high-performance deep learning library. In: WALLACH, H. et al. (Ed.). **Advances in Neural Information Processing Systems**. [S.l.]: Curran Associates, Inc., 2019. p. 8024–8035. DOI: 10.48550/arXiv.1912.01703.

PEREIRA, M. de P. et al. Occurrence of recently diagnosed atrial fibrillation in the immediate postoperative period of myocardial revascularization surgery. although common, a devalued complication. **Revista da Associação Médica Brasileira**, v. 66, n. 11, p. 1473–1475, nov 2020. DOI: 10.1590/1806-9282.66.11.1473.

PROLIFE. **Eletrocardiógrafo Digital HD+**. 2022. Disponível em: <https://prolife.com.br/ecg-digital-hd-touchechg/>. Acesso em: 14 jun. 2023.

PROYECTO ITACA. **The Cardiac Action Potential**. , [s.d.]. Disponível em: <https://www.itaca.edu.es/cardiac-action-potential.htm>. Acesso em: 19 oct. 2023.

RAHMAN, S. et al. Robustness of electrocardiogram signal quality indices. **Journal of The Royal Society Interface**, v. 19, n. 189, abr. 2022. DOI: 10.1098/rsif.2022.0012.

REIFFEL, J. A. Atrial fibrillation: Is rhythm control required, and if so, how, and what is the internist's role? **The American Journal of Medicine**, v. 135, n. 8, p. 939–944, aug 2022. DOI: 10.1016/j.amjmed.2022.03.016.

REYNA, M. A. et al. Issues in the automated classification of multilead ECGs using heterogeneous labels and populations. **Physiological Measurement**, v. 43, n. 8, p. 084001, ago. 2022. DOI: 10.1088/1361-6579/ac79fd.

REYNA, M. A.; SADR, N.; ALDAY, E. A. P. et al. Will two do? varying dimensions in electrocardiography: the physionet/computing in cardiology challenge 2021. In: Computing in Cardiology, 48th., 12-15 Sep 2021, Brno, Czech Republic. **CINC 2021**. [S.l.], v. 48, p. 1–4, 2021.

RIBEIRO, E.; DIAS, F.; SOARES, Q.; KRIEGER, J.; GUTIERREZ, M. Deep learning approach for detection of atrial fibrillation and atrial flutter based on ECG images. In: Anais do XXIII Simpósio Brasileiro de Computação Aplicada à Saúde, 23<sup>o</sup>., 27-30 Jun 2023, São Paulo, Brasil. **SBCAS 2023**. Porto Alegre, RS, Brasil, p. 509–514. ISSN 2763-8952, 2023. DOI: 10.5753/sbcas.2023.229744.

RIBEIRO, E.; SOARES, Q. B.; DIAS, F. M.; KRIEGER, J. E.; GUTIERREZ, M. A. Can deep learning models differentiate atrial fibrillation from atrial flutter? **medRxiv**, 2023. DOI: 10.1101/2023.08.08.23293815.

ROSENBLATT, F. The perceptron: a probabilistic model for information storage and organization in the brain. **Psychological review**, v. 65(6), p. 386–408, nov. 1958. DOI: 10.1037/h0042519.

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. **Nature**, v. 323, n. 6088, p. 533–536, Oct 1986. DOI: 10.1038/323533a0.

SALINET, J. ao L.; SILVA, O. L. Chapter 2 - ECG signal acquisition systems. In: **Developments and Applications for ECG Signal Processing**. [S.l.]: Academic Press, 2019. p. 29–51. ISBN 978-0-12-814035-2. DOI: 10.1016/B978-0-12-814035-2.00008-6.

SATIJA, U.; RAMKUMAR, B.; MANIKANDAN, M. S. A review of signal processing techniques for electrocardiogram signal quality assessment. **IEEE Reviews in Biomedical Engineering**, v. 11, p. 36–52, 2018. DOI: 10.1109/RBME.2018.2810957.

SATTAR, Y.; CHHABRA, L. Electrocardiogram. In: **StatPearls**. Treasure Island (FL): StatPearls Publishing, 2023. Disponível em: <https://www.ncbi.nlm.nih.gov/books/NBK549803/>. Acesso em: 13 oct. 2023.

SAWHNEY, N. S.; FELD, G. K. Diagnosis and management of typical atrial flutter. **Medical Clinics of North America**, v. 92, n. 1, p. 65–85, jan 2008. DOI: 10.1016/j.mcna.2007.08.005.

SAWHNEY, N. S. et al. Diagnosis and management of typical atrial flutter. **Cardiology Clinics**, v. 27, n. 1, p. 55–67, feb 2009. DOI: 10.1016/j.ccl.2008.09.010.

SEIDNER, D. Polyphase antialiasing in resampling of images. **IEEE Transactions on Image Processing**, v. 14, n. 11, p. 1876–1889, 2005. DOI: 10.1109/TIP.2005.854493.

SERVIER MEDICAL ART. **Rib fractures**. Les Laboratoires Servier, 2016. Disponível em: [https://smart.servier.com/smart\\_image/bone-rib-fractures/](https://smart.servier.com/smart_image/bone-rib-fractures/). Acesso em: 24 oct. 2023.

SHI, Y. et al. Robust assessment of ECG signal quality for wearable devices. In: 2019 IEEE International Conference on Healthcare Informatics, 10-13 Jun 2019, ICHI. **ICHI 2019**. [S.l.], p. 1–3, 2019. DOI: 10.1109/ICHI.2019.8904810.

SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. In: International Conference on Learning Representations, 3rd., 7-9 May 2015, San Diego. **ICLR 2015**. [S.l.], 2015. Conference Track Proceedings.

SMITH, L. N. Cyclical learning rates for training neural networks. In: IEEE Winter Conference on Applications of Computer Vision, 24-31 Mar 2017, Santa Rosa, CA, USA. **WACV 2017**. [S.l.], p. 464–472, 2015. DOI: 10.1109/WACV.2017.58.

SOARES, Q.; DIAS, F. M.; RIBEIRO, E.; KRIEGER, J. E.; GUTIERREZ, M. A. Developing a machine learning pipeline for predicting neurological outcomes in comatose cardiac arrest survivors using continuous eeg data. In: Computing in Cardiology, 50th., 1-4 Oct 2023, Atlanta, GA, USA. **CINCC 2023**. [S.l.], v. 50, p. 1–4, 2023. DOI: 10.22489/CinC.2023.165.

SOARES, Q. B.; MONTEIRO, R.; JATENE, F. B.; GUTIERREZ, M. A. A lightweight unidimensional deep learning model for atrial fibrillation detection. In: Computing in Cardiology, 49th., 4-7 Sep 2022, Tampere, Finland. **CINCC 2022**. [S.l.], v. 49, p. 1–4, 2022. DOI: 10.22489/CinC.2022.220.

SÖRNMO, L. et al. Characterization of f waves. In: \_\_\_\_\_. **Atrial Fibrillation from an Engineering Perspective**. Cham: Springer International Publishing, 2018. p. 221–279. ISBN 978-3-319-68515-1. DOI: 10.1007/978-3-319-68515-1\_6.

SRIVASTAVA, N. et al. Dropout: A simple way to prevent neural networks from overfitting. **Journal of Machine Learning Research**, v. 15, n. 56, p. 1929–1958, jan. 2014.

- STEFIL, M.; LIP, G. Y. Atrial fibrillation. **Medicine**, v. 50, n. 8, p. 516–521, jun 2022. DOI: 10.1016/j.mpmed.2022.05.007.
- SUH, J. et al. Learning ECG representations for multi-label classification of cardiac abnormalities. In: *Computing in Cardiology*, 48th., 12-15 Sep 2021, Brno, Czech Republic. **CINC 2021**. [S.l.], v. 48, p. 1–4, 2021. DOI: 10.23919/CinC53138.2021.9662753.
- TAI, C.-T.; CHEN, S.-A. Electrophysiological mechanisms of atrial flutter. **Journal of the Chinese Medical Association**, v. 72, n. 2, p. 60–67, feb 2009. DOI: 10.1016/s1726-4901(09)70024-3.
- TOMPSON, J. et al. Efficient object localization using convolutional networks. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition, Los Alamitos, CA, USA. **CVPR 2015**. [S.l.]: IEEE Computer Society, p. 648–656. ISSN 1063-6919, [s.d.]. DOI: 10.1109/CVPR.2015.7298664.
- UDUPA, A. **Solving 5 challenges of ECG signal acquisition in Wearables**. Texas Instruments, 2021. Disponível em: [https://e2e.ti.com/blogs\\_/b/analogwire/posts/ecg-signal-acquisition-wearables](https://e2e.ti.com/blogs_/b/analogwire/posts/ecg-signal-acquisition-wearables). Acesso em: 02 nov. 22.
- VAIDYANATHAN, P. Multirate digital filters, filter banks, polyphase networks, and applications: a tutorial. **Proceedings of the IEEE**, v. 78, n. 1, p. 56–93, 1990. DOI: 10.1109/5.52200.
- VARRÓ, A. et al. Cardiac transmembrane ion channels and action potentials: cellular physiology and arrhythmogenic behavior. **Physiological Reviews**, v. 101, n. 3, p. 1083–1176, 2021. DOI: 10.1152/physrev.00024.2019.
- VIRTANEN, P. et al. SciPy 1.0: Fundamental algorithms for scientific computing in python. **Nature Methods**, v. 17, p. 261–272, 2020. DOI: 10.1038/s41592-019-0686-2.
- VIVALINK. **Wearable ECG Continuous Heart Monitor - Cardiac Patch - Viva-link**. 2022. Disponível em: <https://www.vivalink.com/wearable-ecg-monitor>. Acesso em: 14 jun. 2023.
- WAGNER, P. et al. PTB-XL, a large publicly available electrocardiography dataset. **Scientific Data**, v. 7, n. 1, may 2020. DOI: 10.1038/s41597-020-0495-6.
- WEBSTER, C. William harvey’s conception of the heart as a pump. **Bulletin of the History of Medicine**, v. 39, n. 6, p. 508–517, 1965. Disponível em: <http://www.jstor.org/stable/44447601>. Acesso em: 14 oct. 2023.
- WERBOS, P. J. **Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences**. Tese (Doutorado) — Committee on Applied Mathematics, Harvard University, Cambridge, MA, 1974.
- WILLCOX, M. E.; COMPTON, S. J.; BARDY, G. H. Continuous ECG monitoring versus mobile telemetry: A comparison of arrhythmia diagnostics in human- versus algorithmic-dependent systems. **Heart Rhythm O2**, v. 2, n. 6, Part A, p. 543–559, 2021. DOI: 10.1016/j.hroo.2021.09.008.

WORLD HEALTH ORGANIZATION. **Cardiovascular diseases (CVDs)**. 2021. Disponível em: [https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)). Acesso em: 12 dez. 2023.

ZHENG, J. et al. A 12-lead electrocardiogram database for arrhythmia research covering more than 10,000 patients. **Scientific Data**, v. 7, n. 1, feb 2020. DOI: 10.1038/s41597-020-0386-x.

\_\_\_\_\_. Optimal multi-stage arrhythmia classification approach. **Scientific Reports**, v. 10, n. 1, feb 2020. DOI: 10.1038/s41598-020-59821-7.

ZIKI, M. A.; LIP, G.; BIKDELI, B. Pulmonary embolism and atrial fibrillation: Two sides of the same coin? a systematic review. **Seminars in Thrombosis and Hemostasis**, v. 43, n. 08, p. 849–863, feb 2017. DOI: 10.1055/s-0036-1598005.

ZIMETBAUM, P.; GOLDMAN, A. Ambulatory arrhythmia monitoring. **Circulation**, v. 122, n. 16, p. 1629–1636, 2010. DOI: 10.1161/CIRCULATIONAHA.109.925610.

# APÊNDICE A – PRODUÇÃO CIENTÍFICA

SOARES, Q.; DIAS, F. M.; RIBEIRO, E.; KRIEGER, J. E.; GUTIERREZ, M. A. Developing a machine learning pipeline for predicting neurological outcomes in comatose cardiac arrest survivors using continuous eeg data. In: Computing in Cardiology, 50th., 1-4 Oct 2023, Atlanta, GA, USA. **CINC 2023**. [S.l.], v. 50, p. 1-4, 2023 <https://doi.org/10.22489/CinC.2023.165>.

SOARES, Q. B.; MONTEIRO, R.; JATENE, F. B.; GUTIERREZ, M. A. A lightweight unidimensional deep learning model for atrial fibrillation detection. In: Computing in Cardiology, 49th., 4-7 Sep 2022, Tampere, Finland. **CINC 2022**. [S.l.], v. 49, p. 1-4, 2022. <https://doi.org/10.22489/CinC.2022.220>.

RIBEIRO, E.; DIAS, F.; SOARES, Q.; KRIEGER, J.; GUTIERREZ, M. Deep learning approach for detection of atrial fibrillation and atrial flutter based on ECG images. In: Anais do XXIII Simpósio Brasileiro de Computação Aplicada à Saúde, 23<sup>o</sup>., 27-30 Jun 2023, São Paulo, Brasil. **SBCAS 2023**. Porto Alegre, RS, Brasil, p. 509–514. ISSN 2763-8952, 2023. <https://doi.org/10.5753/sbcas.2023.229744>.

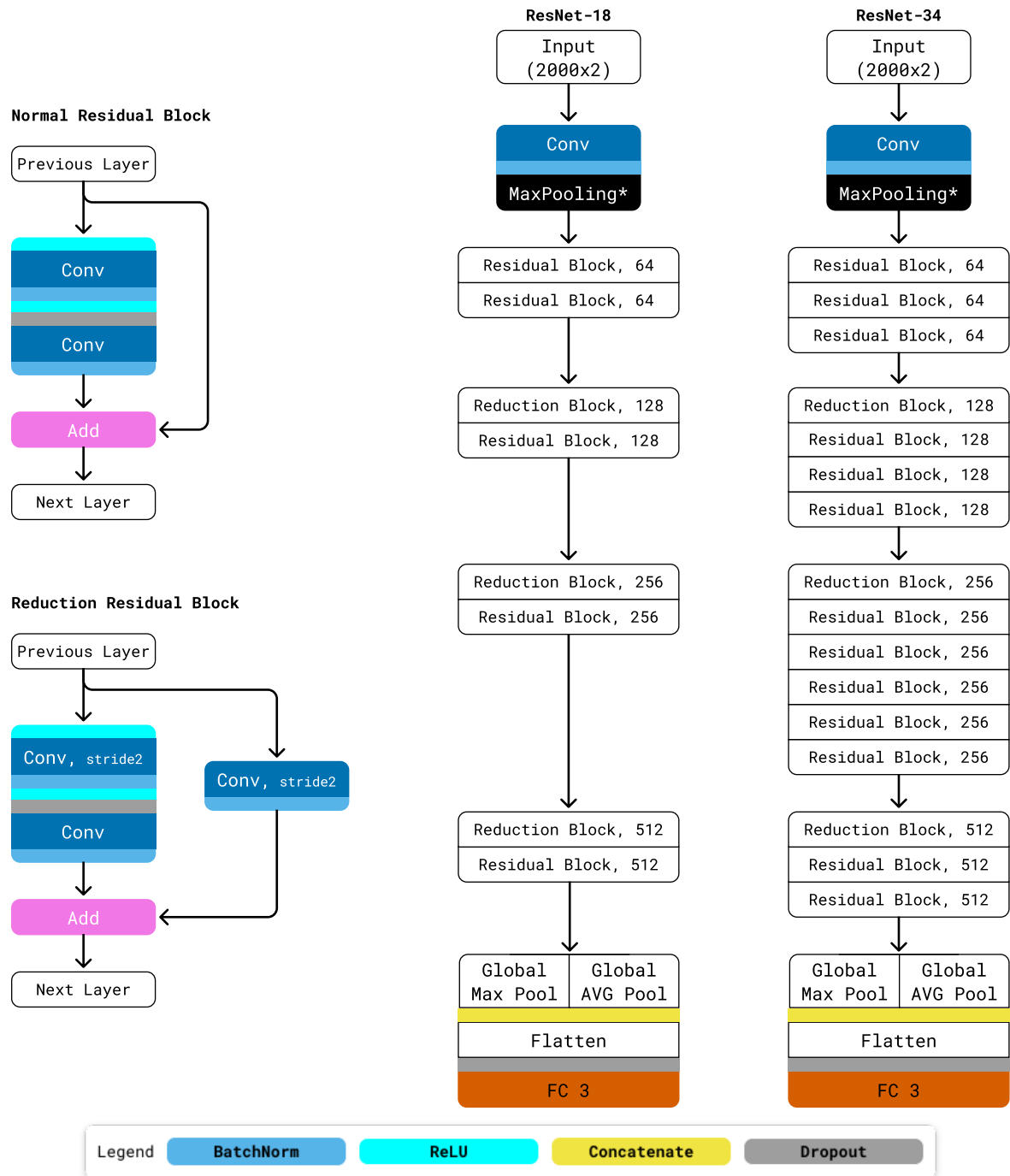
RIBEIRO, E.; SOARES, Q. B.; DIAS, F. M.; KRIEGER, J. E.; GUTIERREZ, M. A. Can deep learning models differentiate atrial fibrillation from atrial flutter? **medRxiv**, 2023. <https://doi.org/10.1101/2023.08.08.23293815>.

DIAS, F. M.; RIBEIRO, E.; SOARES, Q.; KRIEGER, J.; GUTIERREZ, M. A. Predicting age and sex from reduced lead electrocardiograms using deep learning. **TechRxiv**, 2023. <https://doi.org/10.36227/techrxiv.23908089.v1.>>

## APÊNDICE B – RESNET VS LITERESNET

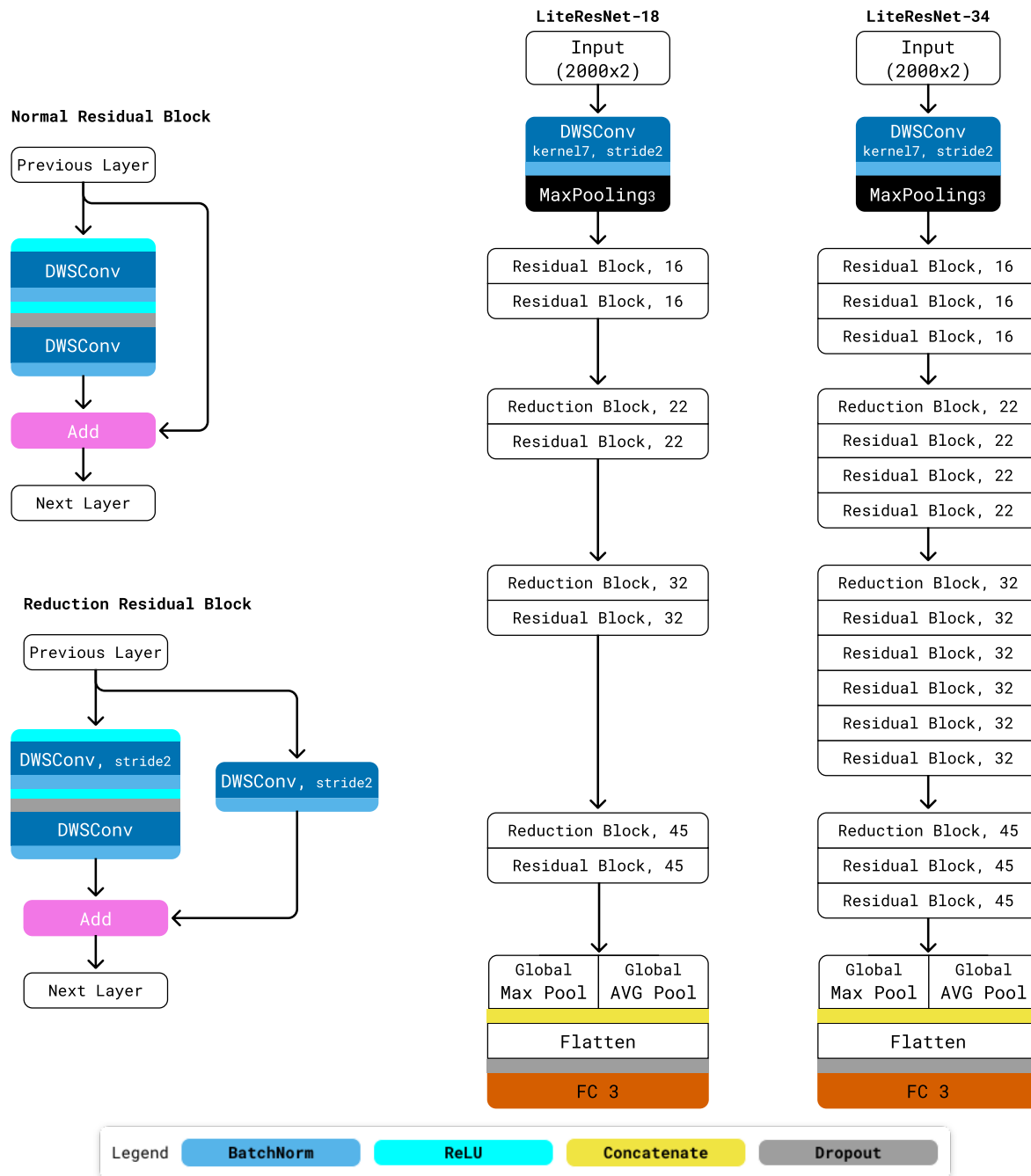
A Figura 42 apresenta a arquitetura dos modelos ResNet-18 e ResNet-34. A arquitetura LiteResNet é obtida aplicando na arquitetura ResNet os passos propostos neste trabalho: substituir camadas convolucionais tradicionais por separáveis e reduzir o número de filtros por camada convolutiva. A Figura 43 apresenta a proposta de arquitetura LiteResNet com 18 e 34 camadas.

Figura 42: Diagrama das arquiteturas ResNet-18 e ResNet-34.



Fonte: De autoria própria.

Figura 43: Diagrama das arquiteturas LiteResNet-18 e LiteResNet-34.



Fonte: De autoria própria.



## **APÊNDICE C – RESULTADOS DOS EXPERIMENTOS PRELIMINARES**

As Tabelas 13, 14 e 15 exibem os resultados obtidos em termos de métricas para a classe FA nos experimentos preliminares de classificação binária, multiclasse e multilabel, respectivamente. As tabelas estão ordenadas em ordem decrescente de F1-score.

Tabela 13: Métricas obtidas para o experimento preliminar de classificação binária.

Modelo	FS	AUC	ACC	PPV	TPR	TNR	FNR	F1-score
<b>LiteVGG-11</b>	128	0,9877	0,9613	0,9635	0,9594	0,9633	0,0406	0,9612
<b>LiteVGG-11</b>	200	0,9848	0,9590	0,9616	0,9562	0,9617	0,0438	0,9589
<b>ResNet-34</b>	200	0,9826	0,9562	0,9566	0,9562	0,9562	0,0438	0,9563
<b>ResNet-18</b>	200	0,9845	0,9547	0,9583	0,9508	0,9586	0,0492	0,9544
<b>ResNet-18</b>	128	0,9836	0,9543	0,9600	0,9484	0,9602	0,0516	0,9540
<b>LiteResNet-18</b>	200	0,9801	0,9539	0,9543	0,9539	0,9539	0,0461	0,9540
<b>ResNet-34</b>	128	0,9847	0,9539	0,9623	0,9453	0,9625	0,0547	0,9535
<b>Hsieh</b>	128	0,9835	0,9488	0,9593	0,9375	0,9601	0,0625	0,9480
<b>Hsieh</b>	200	0,9799	0,9453	0,9481	0,9430	0,9477	0,0570	0,9453
<b>LiteResNet-34</b>	200	0,9795	0,9461	0,9603	0,9312	0,9609	0,0688	0,9453
<b>LiteResNet-34</b>	128	0,9771	0,9418	0,9458	0,9383	0,9453	0,0617	0,9417
<b>LiteResNet-18</b>	128	0,9759	0,9398	0,9546	0,9242	0,9555	0,0758	0,9388
<b>VGG-11</b>	128	0,9733	0,9332	0,9284	0,9398	0,9266	0,0602	0,9337
<b>VGG-19</b>	200	0,9656	0,9285	0,9293	0,9281	0,9289	0,0719	0,9284
<b>Fujita e Cimr</b>	200	0,9579	0,9180	0,9346	0,8992	0,9367	0,1008	0,9164
<b>DWNet</b>	200	0,9638	0,9152	0,9325	0,8961	0,9344	0,1039	0,9131
<b>Fujita e Cimr</b>	128	0,9636	0,9154	0,9398	0,8880	0,9427	0,1120	0,9128
<b>DWNet</b>	128	0,9525	0,8993	0,9042	0,8950	0,9036	0,1050	0,8990
<b>Katsaouni 2</b>	128	0,9440	0,9008	0,9168	0,8828	0,9188	0,1172	0,8979
<b>Katsaouni 4</b>	200	0,9410	0,8887	0,9069	0,8688	0,9086	0,1312	0,8868
<b>Katsaouni 3</b>	128	0,9368	0,8828	0,9055	0,8586	0,9070	0,1414	0,8795
<b>Katsaouni 5</b>	128	0,9122	0,8539	0,8488	0,8742	0,8336	0,1258	0,8592
<b>Katsaouni 4</b>	128	0,9206	0,8629	0,8852	0,8359	0,8898	0,1641	0,8581
<b>Katsaouni 3</b>	200	0,9209	0,8625	0,8935	0,8258	0,8992	0,1742	0,8572
<b>Katsaouni 1</b>	128	0,9077	0,8512	0,8597	0,8438	0,8586	0,1562	0,8499
<b>VGG-11</b>	200	0,9213	0,8906	0,8562	0,8242	0,9570	0,1758	0,8393
<b>VGG-19</b>	128	0,8785	0,8586	0,8153	0,8516	0,8656	0,1484	0,8250
<b>Katsaouni 5</b>	200	0,8775	0,8152	0,8383	0,7883	0,8422	0,2117	0,8096
<b>Katsaouni 2</b>	200	0,8869	0,8246	0,8504	0,7750	0,8742	0,2250	0,8074
<b>Katsaouni 1</b>	200	0,8761	0,8113	0,8125	0,7938	0,8289	0,2062	0,7991
<b>LiteVGG-19</b>	128	0,5000	0,5000	0,2500	0,5000	0,5000	0,5000	0,3333
<b>Nurmaini</b>	128	0,5000	0,5000	0,1667	0,3333	0,6667	0,6667	0,2222
<b>LiteVGG-19</b>	200	0,5000	0,5000	0,1500	0,3000	0,7000	0,7000	0,2000
<b>Nurmaini</b>	200	0,5107	0,5000	0,1500	0,3000	0,7000	0,7000	0,2000

Fonte: De autoria própria.

Tabela 14: Métricas obtidas para a classe FA no experimento preliminar de classificação multiclasse.

Modelo	FS	AUC	Acurácia	Precisão	TPR	TNR	FNR	F1-score
<b>LiteVGG-11</b>	200	0,9839	0,7909	0,9026	0,9359	0,9488	0,0641	0,9185
<b>LiteVGG-11</b>	128	0,9843	0,7865	0,9093	0,9211	0,9535	0,0789	0,9146
<b>LiteResNet-18</b>	200	0,9778	0,7734	0,8826	0,9328	0,9371	0,0672	0,9063
<b>ResNet-18</b>	200	0,9820	0,7846	0,9018	0,9094	0,9492	0,0906	0,9044
<b>ResNet-34</b>	200	0,9809	0,7823	0,9043	0,9008	0,9508	0,0992	0,9012
<b>ResNet-34</b>	128	0,9776	0,7792	0,9047	0,8969	0,9523	0,1031	0,9005
<b>Hsieh</b>	128	0,9718	0,7646	0,8803	0,9219	0,9367	0,0781	0,9002
<b>Hsieh</b>	200	0,9726	0,7555	0,8881	0,9141	0,9414	0,0859	0,8998
<b>ResNet-18</b>	128	0,9785	0,7792	0,9022	0,8922	0,9508	0,1078	0,8963
<b>LiteResNet-34</b>	200	0,9786	0,7680	0,8771	0,9156	0,9355	0,0844	0,8958
<b>LiteResNet-34</b>	128	0,9734	0,7711	0,8761	0,9156	0,9348	0,0844	0,8950
<b>LiteResNet-18</b>	128	0,9736	0,7716	0,8726	0,9117	0,9332	0,0883	0,8911
<b>VGG-11</b>	128	0,9622	0,6984	0,8439	0,9086	0,9152	0,0914	0,8744
<b>Fujita e Cimr</b>	128	0,9468	0,6724	0,8011	0,9039	0,8852	0,0961	0,8481
<b>DWNet</b>	200	0,9593	0,7195	0,8269	0,8719	0,9063	0,1281	0,8470
<b>DWNet</b>	128	0,9335	0,7008	0,7797	0,8555	0,8781	0,1445	0,8152
<b>Katsaouni 2</b>	128	0,9164	0,6232	0,7552	0,8578	0,8586	0,1422	0,8010
<b>VGG-11</b>	200	0,8606	0,6013	0,7193	0,9258	0,7156	0,0742	0,7853
<b>Katsaouni 2</b>	200	0,9011	0,6172	0,7458	0,8227	0,8543	0,1773	0,7791
<b>Katsaouni 3</b>	200	0,9021	0,5987	0,7543	0,8094	0,8719	0,1906	0,7776
<b>Katsaouni 3</b>	128	0,9046	0,6266	0,7484	0,8203	0,8672	0,1797	0,7761
<b>Katsaouni 4</b>	200	0,9016	0,6068	0,7266	0,8227	0,8422	0,1773	0,7700
<b>Katsaouni 1</b>	200	0,8983	0,6065	0,7501	0,7859	0,8687	0,2141	0,7651
<b>Katsaouni 4</b>	128	0,8925	0,6065	0,7241	0,7859	0,8449	0,2141	0,7495
<b>Katsaouni 1</b>	128	0,8718	0,5885	0,7321	0,7297	0,8758	0,2703	0,7140
<b>Katsaouni 5</b>	200	0,8434	0,5609	0,6462	0,7539	0,7891	0,2461	0,6914
<b>Katsaouni 5</b>	128	0,8519	0,5760	0,6613	0,7656	0,8230	0,2344	0,6865
<b>Fujita e Cimr</b>	200	0,8053	0,5466	0,5538	0,5695	0,9250	0,4305	0,5601
<b>VGG-19</b>	200	0,6771	0,4612	0,4678	0,7438	0,5664	0,2562	0,5388
<b>VGG-19</b>	128	0,6798	0,4620	0,3844	0,5703	0,7520	0,4297	0,4420
<b>Nurmaini</b>	128	0,5000	0,3333	0,1667	0,5000	0,5000	0,5000	0,2500
<b>LiteVGG-19</b>	128	0,5000	0,3333	0,1667	0,5000	0,5000	0,5000	0,2500
<b>LiteVGG-19</b>	200	0,5000	0,3333	0,1000	0,3000	0,7000	0,7000	0,1500
<b>Nurmaini</b>	200	0,5000	0,3333	0,0333	0,1000	0,9000	0,9000	0,0500

Fonte: De autoria própria.

Tabela 15: Métricas obtidas para a classe FA no experimento preliminar de classificação multilabel.

Modelo	FS	AUC	Acurácia	Precisão	TPR	TNR	FNR	F1-score
<b>ResNet-34</b>	200	0,9876	0,8223	0,9094	0,9258	0,9762	0,0742	0,9173
<b>LiteVGG-11</b>	128	0,9870	0,8115	0,9063	0,9227	0,9755	0,0773	0,9142
<b>LiteVGG-11</b>	200	0,9859	0,8042	0,9007	0,9203	0,9738	0,0797	0,9101
<b>ResNet-18</b>	200	0,9885	0,8251	0,9204	0,8984	0,9800	0,1016	0,9086
<b>LiteResNet-18</b>	200	0,9855	0,8115	0,9015	0,9164	0,9740	0,0836	0,9082
<b>ResNet-34</b>	128	0,9873	0,8106	0,9201	0,8961	0,9798	0,1039	0,9072
<b>ResNet-18</b>	128	0,9865	0,8126	0,9283	0,8867	0,9822	0,1133	0,9060
<b>LiteResNet-18</b>	128	0,9833	0,8063	0,8886	0,9070	0,9704	0,0930	0,8971
<b>LiteResNet-34</b>	200	0,9857	0,8118	0,8794	0,9133	0,9670	0,0867	0,8950
<b>LiteResNet-34</b>	128	0,9839	0,8049	0,8842	0,8984	0,9696	0,1016	0,8909
<b>Hsieh</b>	128	0,9816	0,7802	0,8767	0,9029	0,9671	0,0971	0,8890
<b>Hsieh</b>	200	0,9795	0,7714	0,9002	0,8698	0,9749	0,1302	0,8833
<b>VGG-19</b>	128	0,9686	0,7322	0,8242	0,8727	0,9521	0,1273	0,8468
<b>DWNet</b>	200	0,9705	0,7319	0,8428	0,8351	0,9594	0,1649	0,8375
<b>VGG-11</b>	128	0,9624	0,7122	0,8167	0,8375	0,9519	0,1625	0,8265
<b>VGG-11</b>	200	0,9626	0,7190	0,8075	0,8281	0,9490	0,1719	0,8169
<b>VGG-19</b>	200	0,9633	0,7087	0,8049	0,8301	0,9477	0,1699	0,8160
<b>Fujita e Cimr</b>	128	0,9649	0,6930	0,8359	0,8002	0,9594	0,1998	0,8139
<b>DWNet</b>	128	0,9549	0,7101	0,8108	0,7645	0,9542	0,2355	0,7861
<b>Fujita e Cimr</b>	200	0,9552	0,6293	0,8089	0,7335	0,9557	0,2665	0,7685
<b>LiteVGG-19</b>	128	0,8890	0,6707	0,7151	0,7281	0,9778	0,2719	0,7214
<b>LiteVGG-19</b>	200	0,8871	0,6771	0,7179	0,7172	0,9790	0,2828	0,7174
<b>Katsaouni 4</b>	200	0,9051	0,5512	0,6933	0,5243	0,9461	0,4757	0,5819
<b>Katsaouni 4</b>	128	0,8841	0,5910	0,5766	0,5167	0,9545	0,4833	0,5308
<b>Katsaouni 5</b>	128	0,8769	0,5202	0,6374	0,3929	0,9560	0,6071	0,4650
<b>Katsaouni 2</b>	200	0,8526	0,5164	0,5072	0,3872	0,9529	0,6128	0,4251
<b>Katsaouni 3</b>	200	0,8331	0,5260	0,5755	0,3655	0,9567	0,6345	0,4220
<b>Katsaouni 1</b>	200	0,8656	0,5055	0,5986	0,3672	0,9490	0,6328	0,4177
<b>Katsaouni 3</b>	128	0,8203	0,5553	0,4666	0,3382	0,9696	0,6618	0,3570
<b>Katsaouni 1</b>	128	0,8304	0,5252	0,4594	0,3225	0,9569	0,6775	0,3476
<b>Katsaouni 5</b>	200	0,8237	0,4667	0,5368	0,2682	0,9676	0,7318	0,3234
<b>Katsaouni 2</b>	128	0,7894	0,4902	0,3463	0,1574	0,9820	0,8426	0,2036
<b>Nurmaini</b>	200	0,5011	0,2035	0,0000	0,0000	1,0000	1,0000	0,0000
<b>Nurmaini</b>	128	0,5000	0,2040	0,0000	0,0000	1,0000	1,0000	0,0000

Fonte: De autoria própria.

# APÊNDICE D – RESULTADOS DO EXPERIMENTO MULTILABEL ROBUSTO

A Tabela 16 apresenta todas as métricas obtidas no experimento robusto de classificação multilabel. Os resultados estão agrupados por classe e, dentro de cada classe, ordenados decrescentemente por F1-score. Para cada uma das classes, o F1-score obtido em cada modelo é comparado por teste t de Student em relação ao modelo com melhor F1-score, sendo apresentado o valor-P obtido deste teste. Modelos que não apresentaram diferença estatisticamente significativa ( $\alpha = 5\%$ ) estão sombreados.

Tabela 16: Métricas obtidas para o experimento robusto de classificação multilabel.

Classe	Modelo	FS	AUC	ACC	PPV	TPR	TNR	FNR	F1-score	p-value
FA	LiteVGG-11	200	0,9918	0,8688	0,9372	0,9458	0,9838	0,0542	0,9414	1,0000
	ResNet-18	200	0,9913	0,8783	0,9579	0,9251	0,9896	0,0749	0,9411	0,7905
	LiteVGG-11	128	0,9920	0,8740	0,9375	0,9442	0,9839	0,0558	0,9408	0,6433
	ResNet-34	200	0,9918	0,8770	0,9555	0,9266	0,9890	0,0734	0,9408	0,5197
	LiteResNet-34	200	0,9915	0,8791	0,9434	0,9375	0,9856	0,0625	0,9404	0,3737

Tabela 16 continuação da página anterior

Classe	Modelo	FS	AUC	ACC	PPV	TPR	TNR	FNR	F1-score	p-value
FA	ResNet-18	128	0,9920	0,8753	0,9515	0,9293	0,9879	0,0707	0,9401	0,3505
	LiteResNet-18	200	0,9914	0,8755	0,9434	0,9357	0,9856	0,0643	0,9394	0,0679
	LiteResNet-18	128	0,9912	0,8721	0,9377	0,9357	0,9841	0,0643	0,9366	0,0005
	Hsieh	200	0,9894	0,8540	0,9319	0,9260	0,9827	0,0740	0,9287	0,0000
	Hsieh	128	0,9884	0,8475	0,9311	0,9221	0,9825	0,0779	0,9264	0,0000
	VGG-11	128	0,9863	0,8337	0,9032	0,9143	0,9749	0,0857	0,9084	0,0000
	VGG-11	200	0,9847	0,8243	0,8911	0,9019	0,9715	0,0981	0,8958	0,0000
RSN	ResNet-18	200	0,9747	0,8783	0,8381	0,8503	0,9580	0,1497	0,8439	1,0000
	LiteResNet-34	200	0,9749	0,8791	0,8370	0,8515	0,9575	0,1485	0,8437	0,9516
	LiteVGG-11	128	0,9756	0,8740	0,8118	0,8717	0,9485	0,1283	0,8406	0,1691
	LiteVGG-11	200	0,9742	0,8688	0,8125	0,8701	0,9488	0,1299	0,8402	0,0831
	LiteResNet-18	200	0,9743	0,8755	0,8271	0,8546	0,9542	0,1454	0,8400	0,0563
	ResNet-34	200	0,9751	0,8770	0,8347	0,8469	0,9569	0,1531	0,8400	0,1035
	ResNet-18	128	0,9736	0,8753	0,8261	0,8515	0,9539	0,1485	0,8378	0,0311
	LiteResNet-18	128	0,9719	0,8721	0,8175	0,8513	0,9513	0,1487	0,8336	0,0002
	Hsieh	200	0,9615	0,8540	0,8094	0,7914	0,9521	0,2086	0,7995	0,0000
	Hsieh	128	0,9549	0,8475	0,8024	0,7687	0,9512	0,2313	0,7838	0,0000
	VGG-11	128	0,9573	0,8337	0,7979	0,7337	0,9514	0,2663	0,7619	0,0000
	VGG-11	200	0,9537	0,8243	0,7954	0,7044	0,9529	0,2956	0,7452	0,0000

Tabela 16 continuação da página anterior

Classe	Modelo	FS	AUC	ACC	PPV	TPR	TNR	FNR	F1-score	p-value
BS	LiteResNet-34	200	0,9951	0,8791	0,9307	0,9629	0,9817	0,0371	0,9465	1,0000
	LiteResNet-18	128	0,9951	0,8721	0,9306	0,9597	0,9818	0,0403	0,9449	0,3078
	ResNet-18	128	0,9952	0,8753	0,9387	0,9512	0,9841	0,0488	0,9448	0,3806
	ResNet-34	200	0,9951	0,8770	0,9349	0,9531	0,9830	0,0469	0,9438	0,0745
	LiteResNet-18	200	0,9950	0,8755	0,9307	0,9570	0,9818	0,0430	0,9435	0,0362
	ResNet-18	200	0,9950	0,8783	0,9363	0,9508	0,9835	0,0492	0,9435	0,0453
BS	LiteVGG-11	128	0,9946	0,8740	0,9192	0,9623	0,9785	0,0377	0,9402	0,0000
	VGG-11	128	0,9913	0,8337	0,9139	0,9559	0,9771	0,0441	0,9344	0,0000
	Hsieh	200	0,9937	0,8540	0,9128	0,9561	0,9767	0,0439	0,9339	0,0000
	VGG-11	200	0,9912	0,8243	0,9129	0,9518	0,9768	0,0482	0,9318	0,0000
	Hsieh	128	0,9933	0,8475	0,9097	0,9542	0,9757	0,0458	0,9312	0,0000
	LiteVGG-11	200	0,9932	0,8688	0,9050	0,9481	0,9746	0,0519	0,9260	0,0000
TS	LiteResNet-34	200	0,9939	0,8791	0,9357	0,9598	0,9832	0,0402	0,9475	1,0000
	ResNet-34	200	0,9937	0,8770	0,9343	0,9606	0,9828	0,0394	0,9472	0,7368
	LiteResNet-18	200	0,9941	0,8755	0,9356	0,9589	0,9832	0,0411	0,9470	0,6141
	ResNet-18	128	0,9936	0,8753	0,9381	0,9557	0,9839	0,0443	0,9468	0,5501
	LiteResNet-18	128	0,9936	0,8721	0,9357	0,9580	0,9832	0,0420	0,9467	0,4824
	LiteVGG-11	128	0,9938	0,8740	0,9328	0,9604	0,9824	0,0396	0,9464	0,3038
	ResNet-18	200	0,9936	0,8783	0,9335	0,9589	0,9826	0,0411	0,9460	0,0717

Tabela 16 continuação da página anterior

Classe	Modelo	FS	AUC	ACC	PPV	TPR	TNR	FNR	F1-score	p-value
TS	LiteVGG-11	200	0,9935	0,8688	0,9308	0,9566	0,9819	0,0434	0,9435	0,0000
	Hsieh	200	0,9918	0,8540	0,9320	0,9522	0,9823	0,0478	0,9419	0,0000
	Hsieh	128	0,9916	0,8475	0,9305	0,9508	0,9819	0,0492	0,9405	0,0000
	VGG-11	128	0,9905	0,8337	0,9322	0,9414	0,9825	0,0586	0,9366	0,0000
	VGG-11	200	0,9902	0,8243	0,9323	0,9411	0,9826	0,0589	0,9366	0,0000
Média	LiteResNet-34	200	0,9889	0,8791	0,9117	0,9279	0,9770	0,0721	0,9195	1,0000
	ResNet-18	200	0,9886	0,8783	0,9165	0,9213	0,9784	0,0787	0,9186	0,2235
	ResNet-34	200	0,9889	0,8770	0,9148	0,9218	0,9779	0,0782	0,9179	0,0531
	LiteResNet-18	200	0,9887	0,8755	0,9092	0,9265	0,9762	0,0735	0,9175	0,0045
	ResNet-18	128	0,9886	0,8753	0,9136	0,9219	0,9775	0,0781	0,9174	0,0630
	LiteVGG-11	128	0,9890	0,8740	0,9003	0,9346	0,9733	0,0654	0,9170	0,0114
	LiteResNet-18	128	0,9880	0,8721	0,9054	0,9262	0,9751	0,0738	0,9154	0,0000
	LiteVGG-11	200	0,9882	0,8688	0,8964	0,9301	0,9723	0,0699	0,9128	0,0000
	Hsieh	200	0,9841	0,8540	0,8965	0,9064	0,9734	0,0936	0,9010	0,0000
	Hsieh	128	0,9821	0,8475	0,8934	0,8989	0,9728	0,1011	0,8954	0,0000
	VGG-11	128	0,9813	0,8337	0,8868	0,8863	0,9715	0,1137	0,8853	0,0000
	VGG-11	200	0,9799	0,8243	0,8829	0,8748	0,9709	0,1252	0,8774	0,0000

Fonte: De autoria própria.



## APÊNDICE E – TRABALHOS FUTUROS

Continuações deste trabalho podem incluir análises de desempenho dos modelos propostos diretamente em arquitetura de hardware embarcado. Estimativas e medições de gasto energético, bem como do tempo de inferência, em microcontroladores poderiam justificar a implantação do monitoramento de ritmo diretamente no hardware do eletrocardiógrafo caso tais gastos sejam inferiores ao consumo da comunicação sem fio com um smartphone, por exemplo. Otimizações específicas para redução do consumo energético também poderiam ser abordadas.

Ainda considerando o uso dos modelos desenvolvidos em plataforma móvel de smartphones, uma possível continuação direta deste trabalho envolveria a otimização dos modelos em termos de quantização dos pesos e poda de nós que pouco contribuem com as saídas dos modelos.

Por fim, caso haja a disponibilidade de grandes bases de dados coletadas com eletrocardiógrafos vestíveis e anotadas quanto a ritmos cardíacos, seria possível avaliar o desempenho dos modelos propostos em um cenário mais realista, onde inclusive o pré-processamento proposto seria colocado à prova.

# ANEXO A – RÓTULOS DO DATASET PHYSIONET CHALLENGE 2021

Um total de 133 diagnósticos estão disponíveis no conjunto de datasets do PhysioNet Challenge 2021, sendo que destes apenas 30 foram considerados para fins de pontuação no desafio. A Tabela 17 lista os diagnósticos considerados para pontuação e o total de exames com cada um destes diagnósticos.

Tabela 17: Diagnósticos pontuados no PhysioNet Challenge 2021.

Abbreviation	Dx	Total
AF	atrial fibrillation	5255
AFL	atrial flutter	8374
BBB	bundle branch block	522
Brady	bradycardia	295
CLBBB	complete left bundle branch block	213
CRBBB	complete right bundle branch block	1779
IAVB	1st degree av block	3534
IRBBB	incomplete right bundle branch block	1857
LAD	left axis deviation	7631
LAnFB	left anterior fascicular block	2186
LBBB	left bundle branch block	1281
LQRSV	low qrs voltages	1599
NSIVCB	nonspecific intraventricular conduction disorder	1768
NSR	sinus rhythm	28971
PAC	premature atrial contraction	3041
PR	pacing rhythm	1481

**Tabela 17 continuação da página anterior**

PRWP	poor R wave Progression	638
PVC	premature ventricular contractions	1279
LPR	prolonged pr interval	392
LQT	prolonged qt interval	1907
QAb	qwave abnormal	2076
RAD	right axis deviation	1280
RBBB	right bundle branch block	3051
SA	sinus arrhythmia	3790
SB	sinus bradycardia	18918
STach	sinus tachycardia	9657
SVPB	supraventricular premature beats	224
TAb	t wave abnormal	11716
TInv	t wave inversion	3989
VPB	ventricular premature beats	659

---