

UNIVERSIDADE FEDERAL DO AMAZONAS
FACULDADE DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

MATEUS DE PAULA DA SILVA

DETECÇÃO DO COMPLEXO QRS EM ELETROCARDIOGRAMAS COM 12
DERIVAÇÕES UTILIZANDO REDES NEURAS CONVOLUCIONAIS.

MANAUS

2023

UNIVERSIDADE FEDERAL DO AMAZONAS
FACULDADE DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

MATEUS DE PAULA DA SILVA

DETECÇÃO DO COMPLEXO QRS EM ELETROCARDIOGRAMAS COM 12
DERIVAÇÕES UTILIZANDO REDES NEURAS CONVOLUCIONAIS.

Dissertação apresentada ao Programa de Mestrado em Engenharia Elétrica da Universidade Federal do Amazonas, como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica na área de concentração Controle e Automação de Sistemas e linha de pesquisa Reconhecimento de Padrões e Otimização.

Orientadora: Profa. Dra. Marly Guimarães Fernandes Costa
Coorientador: Prof. Dr. Cícero Ferreira Fernandes Costa Filho

MANAUS

2023

Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

S586d Silva, Mateus de Paula da
Detecção do complexo QRS em eletrocardiogramas com 12 derivações utilizando redes neurais convolucionais. / Mateus de Paula da Silva . 2023
86 f.: il. color; 31 cm.

Orientadora: Marly Guimarães Fernandes Costa
Coorientador: Cícero Ferreira Fernandes Costa Filho
Dissertação (Mestrado em Engenharia Elétrica) - Universidade Federal do Amazonas.

1. Eletrocardiograma. 2. Complexo QRS. 3. Detecção. 4. 12 derivações. 5. Redes Neurais Convolucionais. I. Costa, Marly Guimarães Fernandes. II. Universidade Federal do Amazonas III. Título

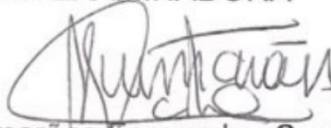
MATEUS DE PAULA DA SILVA

**DETECÇÃO DO COMPLEXO QRS EM ELETROCARDIOGRAMAS
COM 12 DERIVAÇÕES UTILIZANDO REDES NEURAIS
CONVOLUCIONAIS**

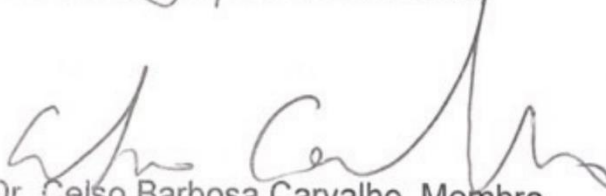
Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas, como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica na área de concentração Controle e Automação de Sistemas.

Aprovada em 19 de maio de 2023.

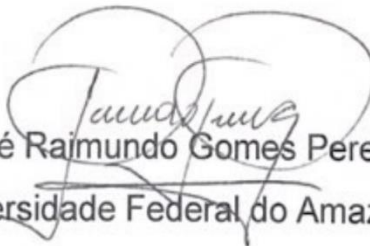
BANCA EXAMINADORA



Prof^a. Dra. Marly Guimarães Fernandes Costa, Presidente
Universidade Federal do Amazonas



Prof. Dr. Celso Barbosa Carvalho, Membro
Universidade Federal do Amazonas



Prof. Dr. José Raimundo Gomes Pereira, Membro
Universidade Federal do Amazonas

AGRADECIMENTOS

Agradeço aos meus orientadores, Prof^a. Dr^a. Marly Guimarães Fernandes Costa e Prof. Dr. Cícero Ferreira Fernandes Costa Filho, pela orientação dada, pelo aprendizado adquirido e, principalmente, pela confiança em mim depositada.

Gostaria de fazer um agradecimento especial à minha falecida mãe, Oneida Pereira de Paula, a quem dedico este trabalho. Ela sempre foi uma grande inspiração e motivadora em minha vida e tenho certeza de que estaria orgulhosa de ver esta realização.

Agradeço a todos os amigos e familiares que me apoiaram durante essa jornada, pelo apoio desde o início da minha vida estudantil e em todas as fases da minha vida.

A Universidade Federal do Amazonas e, em especial, ao Centro de Tecnologia Eletrônica e da Informação – CETELI – pelo provimento de toda infraestrutura para realização deste trabalho.

A CAPES, pela concessão de bolsa que permitiu dedicação exclusiva aos estudos.

Por fim, espero que esta dissertação possa contribuir para o avanço do conhecimento na área e, de alguma forma, beneficiar a sociedade.

RESUMO

As doenças cardiovasculares (CDVs) são as principais causas de mortes no mundo. De acordo com a Organização Mundial da Saúde (OMS) cerca de 17,9 milhões de pessoas morreram de CDVs em 2019, sendo a maior ocorrência em países de baixa e média renda. As CDVs têm consequências econômicas significativas que afetam indivíduos, sistemas de saúde e a sociedade. A detecção precoce é a principal forma de prevenção. O Eletrocardiograma (ECG) é a técnica amplamente utilizada para detecção de CDVs. Ele representa a atividade elétrica do coração, este possui várias ondas sendo a mais proeminente o complexo QRS. Os sinais do ECG são captados em várias regiões do corpo, sendo o ECG de 12 derivações comumente utilizado nas práticas clínicas. O uso de técnicas de aprendizado de máquina e, nos últimos tempos o aprendizado profundo, tem sido foco de muitas pesquisas com o fim de detecção do complexo QRS. Entretanto a pesquisa em base de dados literárias evidenciou a carência desses estudos aplicados individualmente às 12 derivações do ECG. Neste estudo, propõe-se avançar na detecção do complexo QRS em todas as 12 derivações do ECG. Para tal, usamos o banco de dados público, o INCART, no qual realizamos um processo de rotulagem dos intervalos QRS de todos os sinais disponibilizados no referido banco. A metodologia investiga três arquiteturas de redes neurais convolucionais (CNN1, 2 e 3) com diferentes tamanhos de campos receptíveis. A arquitetura que apresentou melhor desempenho médio das 12 derivações foi a CNN2, com uma precisão de 76,44%, *recall* de 99,90% e *F1-score* de 86,50%. Os valores de precisão e *F1-score* foram afetados pelo fato do modelo ter apresentado falsos positivos. Estes foram minimizados por um passo adicional de pós-processamento. Com o pós-processamento, as métricas de desempenho passaram a ser: precisão de 99,98%, *recall* de 99,90% e *F1-score* de 99,94%, superando o modelo de CNN apresentado por Xiang et. al. (2018) na mesma base de dados.

Palavras-chave: Eletrocardiograma, ECG, Complexo QRS, Detecção, 12 derivações, Redes Neurais Convolucionais.

ABSTRACT

Cardiovascular diseases (CVDs) are the leading cause of death in the world. According to the World Health Organization (WHO) about 17.9 million people died from CVDs in 2019, with the highest occurrence in low- and middle-income countries. CVDs have significant economic consequences that affect individuals, healthcare systems, and society. Early detection is the main form of prevention. The Electrocardiogram (ECG) is the widely used technique for detecting CVDs. It represents the heart's electrical activity, and has several waves, the most prominent of which is the QRS complex. The ECG signals are captured in various regions of the body, and the 12-lead ECG is commonly used in clinical practice. The use of machine learning techniques, and in recent times deep machine learning, has been the focus of much research in order to detect the QRS complex. However, a literature database search evidenced the lack of such studies applied individually to the 12 leads of the ECG. In this study, we propose to advance the detection of the QRS complex in all 12 ECG leads. To this end, we use the public dataset, INCART, in which we label the QRS intervals of all the signals available in this dataset. The methodology investigates three convolutional neural network architectures (CNN1, 2 and 3) with different sizes of receptive fields. The architecture that showed the best average performance for the 12 leads-ECG was CNN2, with an accuracy of 76.44%, recall of 99.90%, and F1-score of 86.50%. The accuracy and recall values were affected by the model's false positives. These were minimized by an additional post-processing step. With this post-processing, the performance metrics became: accuracy 99,98%, recall de 99,90% and F1-score, outperforming the CNN model presented by Xiang et. al. (2018) on the same dataset.

Keywords: Electrocardiogram, ECG, QRS Complex, Detection, 12 leads, Convolutional Neural Networks.

LISTA DE FIGURAS

Figura 1: O complexo ECG com as segmentações e intervalos definidos no sinal...	19
Figura 2: Representação do intervalo QRS em um sinal ECG.....	34
Figura 3: Diferentes apresentações do complexo QRS.	35
Figura 4: Rede Neural Artificial totalmente conectada (densa).	36
Figura 5: Arquitetura de uma CNN contendo as principais camadas de processamento.	39
Figura 6: Ilustração da operação de convolução linear.	41
Figura 7: Exemplo da operação matemática na camada de pooling.	44
Figura 8: Modelo de Operação da convolução transposta.	45
Figura 9: Exemplo do efeito nas conexões dos neurônios com e sem o uso de dropout.	47
Figura 10: Comparação do ADAM com outros métodos de otimização. Fonte: (KINGMA; BA, 2014)	48
Figura 11: Exemplo de uma CNN para a segmentação semântica.....	49
Figura 12: Diagrama em blocos apresentando as etapas da metodologia utilizada nesse trabalho para detecção do complexo QRS.	50
Figura 13: Registro I01 do paciente 1 com as 12 derivações do ECG obtido da base de dados INCART (amplitude 5 mm/mV e velocidade de 25 mm/segundos). Na parte superior da figura observa-se as anotações feitas para complexo QRS.....	53
Figura 14: Ilustração da amostra 1 extraída do registro I01 do paciente 1, com identificação dos pontos Q, R e S no sinal da derivação II.	54
Figura 15: (a) Imagem de uma amostra de sinal extraída do sinal ECG conforme procedimento descrito no item 4.1.3; (b) padrão ouro correspondente a amostra apresentada em (a).	56
Figura 16: Representação da Arquitetura CNN1 de rede neural convolucional.	59
Figura 17: Representação das Arquiteturas de CNN 2 e 3 de rede neural convolucional.....	60
Figura 18: Divisão do conjunto de dados para ser utilizado nos experimentos com a rede neural convolucional, com 80% dos dados no conjunto de treinamento e 20% dos dados no conjunto de teste.	61
Figura 19: Procedimento de treinamento para as 12 derivações do ECG.	63

Figura 20: Passo 1 da etapa de pós-processamento: delimitação no sinal original dos intervalos <i>QS</i> identificados pela rede neural.	64
Figura 21: Passo 2 da etapa de pós-processamento: identificação dos pontos de máximos e mínimos para o cálculo da amplitude do sinal em cada intervalo <i>QS</i> detectado.....	64
Figura 22: Gráfico de convergência das curvas de acurácia e perda em função do número de iterações do treinamento, para a arquitetura CNN1. Foram utilizados dados da derivação I.....	68
Figura 23: Resultado antes e depois do pós-processamento na amostra do ECG do paciente 20: (a) sinal original de ECG com o padrão ouro sobreposto; (b) intervalos <i>QS</i> detectados pela rede convolucional; (c) intervalos <i>QS</i> detectados após a aplicação do pós-processamento.	72
Figura 24: Resultado antes e depois do pós-processamento nas amostras do paciente 21. (a) Sinal original de ECG com o padrão ouro sobreposto; (b) intervalos <i>QS</i> detectados pela rede convolucional; (c) intervalos <i>QS</i> detectados após a aplicação do pós-processamento.	73

LISTA DE TABELAS

Tabela 1: Relação das informações dos pacientes presentes na base de dados INCART.....	52
Tabela 2: Filtros por unidade convolutiva para as CNN 1, 2 e 3.	61
Tabela 3: Representação dos parâmetros para o treinamento do modelo.....	62
Tabela 4: Os resultados das métricas de precisão, recall e F1-score para a arquitetura CNN1, 2 e 3, no conjunto de teste, para as 12 derivações, antes e depois do pós-processamento.....	70
Tabela 5: Análise comparativa dos trabalhos com propostas semelhantes utilizando base de dados com 12 derivações.....	76
Tabela 6: Análise comparativa entre o estado da arte e o método proposto, comparação entre abordagens que utilizam modelos baseados em redes neurais convolutivas.....	78

LISTA DE QUADROS

Quadro 1 – Resumo da análise realizada nos trabalhos que tratam da detecção do complexo QRS em ECG.	25
--	----

LISTA DE SIGLAS

ADAM	<i>Adaptive Moment Estimation</i>
CNN	<i>Convolutional Neural Network</i>
SVM	<i>Support Vector Machine</i>
ECG	Eletrocardiograma
PDS	Processamento Digital de Sinais
ML	<i>Machine Learning</i>
DL	<i>Deep Learning</i>
CSE	<i>Common Standards for Electrocardiography</i>
INCART	<i>Institute of Cardiological Technics</i>
EDB	<i>European ST-T Database</i>
PCDB	<i>PhysioNet Computing DataBase</i>
CPSC	<i>China Physiological Signal Challenge</i>
NSTDB	<i>Noise Stress Test DataBase</i>
LUDB	<i>Lobachevsky University Electrocardiography Database</i>
FIR	<i>Finite Impulse Response</i>
RNA	Redes Neurais Artificiais
CDVs	Doenças Cardiovasculares
ReLU	<i>Rectified Linear Unit</i>
VMD	<i>Variational Mode Decomposition</i>
LSTM	<i>Long Short-Term Memory Networks</i>
SGDM	<i>Stochastic Gradient Descent With Momentum</i>
RMSProp	<i>Root Mean Square Propagation</i>

SUMÁRIO

1	INTRODUÇÃO	14
1.1	Objetivo Geral	16
1.2	Objetivos Específicos	16
1.3	Organização do Trabalho	16
2	REVISÃO DA LITERATURA	18
2.1	Introdução	18
2.2	Análise dos trabalhos relacionados	19
2.2.1	Trabalhos utilizando Processamento Digital de Sinais	20
2.2.2	Trabalhos utilizando PDS e Aprendizado de Máquina	21
2.2.3	Trabalhos utilizando Aprendizado de Máquina	22
2.3	Discussão dos trabalhos	30
3	REFERENCIAL TEÓRICO	32
3.1	O Eletrocardiograma	32
3.1.1	Derivações do ECG	33
3.1.2	Complexo QRS	34
3.2	Redes Neurais Artificiais (RNA)	35
3.2.1	Funções de Ativações	37
3.2.2	Treinamento da rede	37
3.3	Redes Neurais Convolucionais	38
3.3.1	Camada Convolutiva	40
3.3.2	Camada <i>Batch Normalization</i>	41
3.3.3	Camada de Unidades Lineares Retificadoras (ReLU)	42
3.3.4	Camada de <i>Pooling</i> (ou sub-amostragem)	43
3.3.5	Camada de Convolução Transposta (ou sobre-amostragem)	44
3.3.6	Camada de Classificação	45
3.3.7	Camada <i>Dropout</i>	46

3.4	Processo de Treinamento de uma Rede Convolutiva	47
3.4.1	Método de Otimização	47
3.4.2	Segmentação Semântica.....	48
4	MATERIAIS E MÉTODOS	50
4.1	Materiais.....	51
4.1.1	Definição do ambiente de trabalho	51
4.1.2	Banco de Dados	51
4.1.3	Modelagem da base de dados	53
4.1.4	Geração do Sinal Padrão Ouro (<i>Ground Truth</i>).....	55
4.2	Métodos	56
4.2.1	Pré-Processamento dos dados.....	56
4.2.2	Definição das Arquiteturas das Redes Neurais Convolucionais	57
4.2.3	Experimentos.....	61
4.2.4	Pós-Processamento	63
4.2.5	Métricas de Desempenho.....	65
5	RESULTADOS E DISCUSSÕES.....	68
5.1	Convergência da Rede Neural Convolutiva.....	68
5.2	Resultados de Testes dos Modelos	69
5.3	Discussão dos resultados	74
6	CONCLUSÕES	80
6.1	Trabalhos Futuros	80
7	REFERÊNCIAS.....	82

1 INTRODUÇÃO

De acordo com a Organização Mundial de Saúde (OMS), as doenças cardiovasculares (CVDs, do inglês *cardiovascular diseases*) são as principais causas de morte no mundo. Estima-se que 17,9 milhões de pessoas morreram de CVDs em 2019, representando 32% de todas as mortes globais. Mais de 75% dessas mortes ocorreram em países de baixa e média renda. Dos 17 milhões de mortes prematuras (com menos de 70 anos) por doenças não transmissíveis em 2019, 38% foram causadas por CVDs (*World Health Organization, 2022*)

As consequências econômicas das doenças cardiovasculares afetam não apenas os indivíduos, mas os sistemas de saúde e a sociedade. As estimativas da *World Health Organization (2022)* mostram que os custos globais com CVDs em 2010 foram de US\$ 863 bilhões de dólares e em 2030 aumentará para US\$ 1044 bilhões de dólares.

A maioria das doenças cardiovasculares pode ser prevenida, logo é importante detectar o mais precocemente possível e seguir as orientações médicas.

O diagnóstico de pacientes com CVD, graças a tecnologias como o Eletrocardiograma (ECG), tem sido promissor (TIMMIS *et al.*, 2020). O ECG é amplamente utilizado como a principal técnica para a detecção de doenças cardiovasculares, como infarto do miocárdio, hipertrofia ventricular esquerda, doença arterial coronariana, entre outras. Isso se deve ao fato de que o ECG é capaz de registrar a atividade elétrica do coração, fornecendo informações valiosas sobre a saúde do músculo cardíaco (GOLDBERGER, 2018), permitindo aos médicos identificar e avaliar padrões anormais que possam indicar uma doença em desenvolvimento.

O sinal ECG é composto por diferentes ondas (P, Q, R, S e T) que representam diferentes fases do ciclo cardíaco, sendo o complexo QRS a mais importante visão do ECG, pois sua correta identificação representa a despolarização ventricular, o que é crucial na detecção de várias doenças cardiovasculares, incluindo arritmia ventricular.

As derivações do ECG são representações gráficas das diferenças de potencial elétrico entre pares de eletrodos colocados em locais específicos no corpo do paciente, que permitem a visualização da atividade elétrica do coração em diferentes planos e perspectivas.

Na visão de Goovaerts *et al.* (2018) a detecção do complexo QRS nas 12 derivações do ECG é um fator importante na identificação de doenças cardiovasculares, pois permite a obtenção de informações completas sobre a atividade elétrica do coração. A detecção do complexo QRS no ECG é amplamente estudada na literatura. O uso de algoritmos de aprendizado de máquina, especialmente redes neurais convolucionais, na última década, tem sido uma abordagem comum para a detecção do complexo QRS (NAGAL e SHARMA, 2014; ZHONG *et al.*, 2018; SILVA *et al.*, 2018; YUEN *et al.*, 2019; LIU *et al.*, 2020; HABIB *et al.*, 2021). No entanto, não há estudos que aplicam essa técnica nas 12 derivações do ECG individualmente e que utilizem uma base de dados de sinais de ECG rotulada com os intervalos de início e término do segmento \overline{QS} em um modelo de segmentação semântica. A análise simultânea das 12 derivações pode fornecer informações valiosas para a detecção precisa do complexo QRS e, portanto, é uma área que pode ser explorada em futuras pesquisas (Liu *et al.*, 2021). Além disso, o uso de uma base de dados rotulada, com o intervalo do complexo QRS anotado, permitiria uma maior precisão e personalização na detecção do complexo QRS.

Nesta pesquisa, a proposta é contribuir com o avanço do estado da arte do tema de detecção do complexo QRS nas 12 derivações do ECG. Para isso, foi utilizada a base de dados INCART (GOLDBERGER *et al.*, 2000) de 12 derivações a qual foi processada para fins de rotulação. A metodologia utilizada nesta pesquisa para obter o intervalo de início e término do complexo QRS, baseou-se na utilização de redes neurais convolucionais. A rede neural que utilizamos implementa a técnica de segmentação semântica para a detecção do complexo QRS.

Além disso, foi desenvolvido um método eficiente de pós-processamento, que possibilitou elevar o desempenho de precisão, *recall* e *f1-score* na tarefa de detecção para o estado da arte.

Desta forma, contribuimos para o avanço do estado da arte na detecção do complexo QRS nas 12 derivações do ECG, fornecendo uma metodologia eficiente para a detecção deste sinal, assim como uma base de dados rotulada para futuras pesquisas no tema.

1.1 Objetivo Geral

Propor uma nova metodologia para a detecção do complexo QRS em eletrocardiogramas de 12 derivações baseada em redes neurais convolucionais.

1.2 Objetivos Específicos

- a) Avaliar diferentes arquiteturas de redes convolucionais na detecção do complexo QRS;
- b) Comparar o desempenho das arquiteturas propostas com o estado da arte.
- c) Propor um método de pós-processamento, associado às redes convolucionais, que melhore o desempenho da detecção do complexo QRS.

1.3 Organização do Trabalho

Este trabalho está organizado conforme os tópicos mostrados a seguir:

- Capítulo 1: Introdução;
- Capítulo 2: Revisão da literatura;
- Capítulo 3: Referencial teórico;
- Capítulo 4: Materiais e métodos;
- Capítulo 5: Resultados preliminares e discussões.
- Capítulo 6: Conclusões
- Capítulo 7: Referências

O Capítulo 1 contextualiza o problema do ECG e a detecção do complexo QRS nas 12 derivações, mostra o objetivo geral e os objetivos específicos da pesquisa.

No Capítulo 2 há uma revisão da literatura disponível sobre o tema, efetua uma análise dos trabalhos publicados, tanto os que utilizam processamento digital de sinais quanto os que utilizam aprendizado de máquina para a detecção do complexo QRS. O capítulo engloba também uma discussão dos trabalhos encontrados.

No Capítulo 3 apresenta os conceitos teóricos relacionados ao eletrocardiograma, incluindo as derivações do ECG e o complexo QRS. Além disso,

são apresentados conceitos sobre redes neurais artificiais, incluindo redes neurais convolucionais, suas camadas e o processo de treinamento.

No Capítulo 4 descreve os materiais utilizados na pesquisa, incluindo o ambiente de trabalho, o banco de dados e a modelagem da base de dados. Além disso, são descritos os métodos utilizados na pesquisa, incluindo o pré-processamento dos dados, a definição da arquitetura da rede neural, os experimentos e o pós-processamento.

O Capítulo 5 apresenta os resultados preliminares da simulação da rede neural, incluindo gráficos mostrando a convergência da rede durante o treinamento e os resultados dos testes do modelo. Além disso, são realizadas discussões sobre os resultados até aqui obtidos.

Por fim, no Capítulo 6 há as conclusões da pesquisa e propostas futuras, no Capítulo 7 as referências com base na literatura que fundamentam a pesquisa.

2 REVISÃO DA LITERATURA

2.1 Introdução

Neste capítulo apresenta-se a revisão da literatura sobre o tema de detecção do complexo QRS em eletrocardiogramas (ECG), com o propósito de conhecer o estado da arte sobre este tema. As pesquisas foram realizadas nas bases literárias: *IEEE Xplore*, *PubMed*, *Web of Science* e *Engineering Village*, acessadas a partir do Portal de Periódicos - CAPES. Os artigos foram obtidos utilizando a estratégia de busca avançada, com os termos de busca apropriados.

Buscou-se principalmente por artigos de revista, por apresentarem maior credibilidade e confiabilidade. Quanto ao período, a busca, foi feita considerando dois intervalos. O primeiro compreende artigos mais recentes, publicados nos últimos 5 anos (2018 a 2022), e o segundo, artigos publicados antes de 2017. Isto possibilita acompanhar a evolução do tema e as abordagens adotadas pelos autores ao longo dos anos.

O batimento cardíaco é um dos sinais vitais mais básicos e importantes considerados na medicina. Portanto, os parâmetros que descrevem os batimentos cardíacos são uma ferramenta importante para monitorar a condição cardiológica de um paciente. Um desses parâmetros é a Frequência Cardíaca (FC), que descreve o ritmo cardíaco. O cálculo da FC é feito por meio do monitoramento da atividade elétrica do coração, através do método denominado de eletrocardiografia.

A eletrocardiografia produz o eletrocardiograma (ECG), um sinal biomédico que mostra a atividade elétrica do coração. Trata-se de uma ferramenta de representação gráfica, que registra a atividade elétrica do coração, mais especificamente, constitui-se numa representação gráfica das diferenças de potenciais geradas. Basicamente, o ECG registra dois fenômenos, quais sejam: a despolarização, que é a transmissão do estímulo através do músculo cardíaco, e a repolarização, que é o retorno do músculo cardíaco estimulado ao estado de repouso.

Habitualmente, para o registro do ECG são utilizadas 12 derivações, que são pares de terminais de polaridade definida, conectados a eletrodos exploradores (MOFFA e SANCHES, 2010). O ECG envolve o registro das principais deflexões (P, QRS, T e U) nas 12 derivações. Mais detalhes são apresentados no capítulo de

fundamentação teórica. A Figura 1 apresenta a forma de onda do eletrocardiograma normal. A onda P é o primeiro registro identificável no ECG, que representa a despolarização dos átrios. O segundo registro identificável e mais marcante no ECG é o complexo QRS, que representa a despolarização (contração) ventricular. A onda T corresponde à repolarização ventricular. A repolarização dos átrios é superposta pelo complexo QRS.

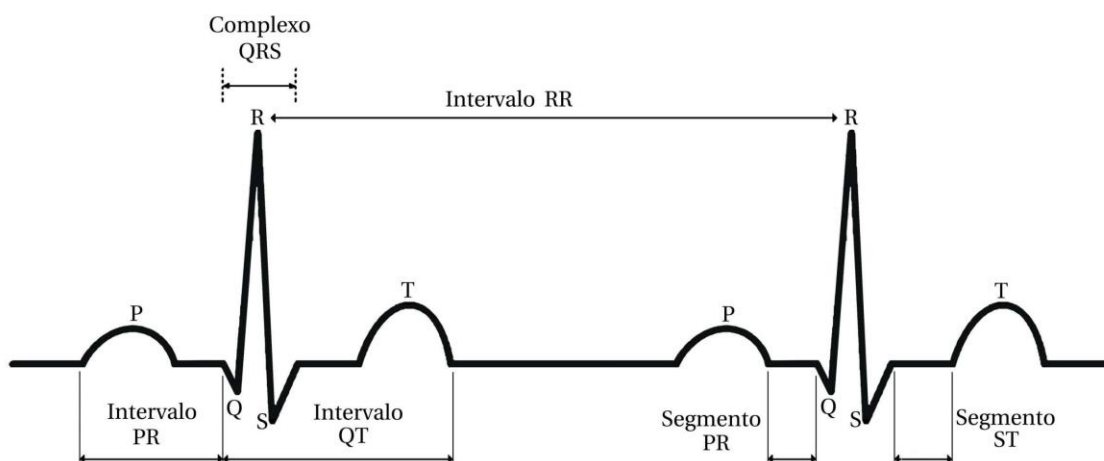


Figura 1: O complexo ECG com as segmentações e intervalos definidos no sinal.
Fonte: Adaptado de Goldberger (2018).

2.2 Análise dos trabalhos relacionados

O levantamento bibliográfico das publicações que se debruçam no estudo do complexo QRS mostra três abordagens/fins de pesquisa. Elas se dividem em: classificação do ECG, detecção do complexo QRS e segmentação do complexo QRS. Ressalta-se que os artigos que utilizam algoritmos com o propósito de classificação do ECG como normal ou associado a uma dada doença (arritmia) não estão relacionados com o objetivo desta pesquisa. Logo, apenas as abordagens de segmentação e principalmente detecção serão aqui discutidas.

A tarefa de segmentação do ECG, visa a identificação/extração das ondas e/ou intervalos constituintes desse sinal bioelétrico, entre os quais o intervalo ou complexo QRS. Tal tarefa é observada nos trabalhos de (BORDE *et al.*, 2019; LIU *et al.*, 2020 e DURAJ *et al.*, 2022).

Com respeito a abordagem de detecção, a tarefa é mais específica, pois visa o complexo QRS apenas. Os artigos mais antigos que tratam dessa tarefa utilizam

exclusivamente técnicas de Processamento Digital de Sinais (PDS). Podemos exemplificar essa afirmação com os seguintes trabalhos (DE OLIVEIRA *et al.*, 2004; HUANG e WANG, 2009; JALIL *et al.*, 2010; LUTFULLAH *et al.*, 2011; HU *et al.*, 2014; PANDIT *et al.*, 2017; LU *et al.*, 2018 e ZHAO *et al.*, 2018). Há também trabalhos que utilizam Aprendizado de Máquina, ML (do inglês *Machine Learning*). Como exemplos destes, citamos (NAGAL e SHARMA, 2014; ZHONG *et al.*, 2018; SILVA *et al.*, 2018; YUEN *et al.*, 2019; LIU *et al.*, 2020; HABIB *et al.*, 2021 e etc.), bem como os que combinam as duas abordagens (MEHTA e LINGAYAT, 2008 e GOOVAERTS *et al.*, 2019). O Quadro 1, possibilita visualizar a evolução das técnicas de detecção do complexo QRS em sinais de eletrocardiogramas (ECG).

2.2.1 Trabalhos utilizando Processamento Digital de Sinais

Os detectores do complexo QRS propostos inicialmente utilizaram abordagens baseadas em Processamento Digital de Sinais (PDS). Esses trabalhos (PAN e TOMPKINS, 1985 e LI *et al.*, 1995) utilizavam: filtro passa banda, derivadas, funções quadráticas, regras de limiar e transformada *wavelet* para a detecção do complexo QRS.

O complexo QRS em um sinal ECG normal é a região mais proeminente quando comparado aos demais intervalos, conforme pode ser observado na Figura 1. Baseado nesse fato, os autores (DE OLIVEIRA e CORTEZ, 2004; HUANG e WANG, 2009; JALIL *et al.*, 2010; LUTFULLAH *et al.*, 2011; PANDIT *et al.*, 2017; LU *et al.*, 2018 e ZHAO *et al.*, 2018) utilizam técnicas de limiar para separar esta região das demais.

A saliência da onda R no complexo QRS pode não ser aparente na presença de ruídos, patologias e em derivações diferentes. Por isso, a maioria das abordagens utilizam uma etapa de pré-processamento, tais como: filtros que removem a interferência da rede e/ou ruídos de alta frequência; remoção da linha de base; normalização do sinal.

Os detectores mais comuns fazem uso da transformada *wavelet* (DE OLIVEIRA e CORTEZ, 2004; HUANG e WANG, 2009; JALIL *et al.*, 2010). Outros utilizam análise de componentes principais (*Principal Component Analysis - PCA*) (HUANG e WANG, 2009), janela de Chebyshev (LUTFULLAH *et al.*, 2011), algoritmo de morfologia do triângulo modificado (HU *et al.*, 2014), máxima e mínima diferença (PANDIT *et al.*,

2017). Em alguns trabalhos ocorre a combinação de duas ou mais técnicas. Devido à diferença de amplitude entre as ondas R, P e/ou T, a maioria dos detectores, exceto o proposto por Hu *et al.* (2014), fazem uso do limiar (de localização e/ou de amplitude) para evitar falsos positivos.

Os algoritmos de detecção são testados em uma única derivação do ECG (DE OLIVEIRA e CORTEZ, 2004; JALIL *et al.*, 2010; HU *et al.*, 2014; PANDIT *et al.*, 2017 e LU *et al.*, 2018) ou nas 12 derivações (HUANG e WANG, 2009; LUTFULLAH *et al.*, 2011 e ZHAO *et al.*, 2018) em sinais de ECG normais ou com patologias.

Utilizando técnicas de PDS, os resultados da detecção do complexo QRS apresentaram altas taxas de desempenho em sinais de ECG de derivação única e em sinais de ECG com 12 derivações. Os melhores valores de *recall* (sensibilidade) e precisão foram obtidos por Huang e Wang (2009), 99,98% e 99,99%, respectivamente. A base de dados utilizada pelo autor foi do Instituto de Técnicas Cardiológicas de São Petersburgo, conhecida por INCART, contendo 75 amostras de ECG com 12 derivações.

No entanto, é observado o baixo desempenho dos detectores na presença de patologias, resultando em altas taxas de falsos positivos. Assim, para amenizar esse problema é necessário criar várias regras de decisões para cada situação, o que eleva o custo computacional, tornando inviável a aplicação para sistemas embarcados.

2.2.2 Trabalhos utilizando PDS e Aprendizado de Máquina

Abordagens híbridas (PDS e ML) de detecção do complexo QRS são identificadas em dois trabalhos (MEHTA e LINGAYAT, 2008; GEOVAENTS *et al.*, 2019).

O trabalho de Mehta e Lingayat (2008) utiliza uma medida estatística de entropia para detectar a presença ou não do complexo QRS (região de QRS e região de não QRS), através do classificador Máquina de Vetores de Suporte (MVS). A ideia do detector é aplicada a sinais de ECG com derivação única e com 12 derivações. A base de dados utilizada foi a *Common Standards for Electrocardiography* (CSE). Os autores utilizaram uma etapa de pré-processamento com filtros que removem a linha de base e interferências de rede. As métricas apresentadas por Mehta e Lingayat

(2008) mostram o melhor desempenho com 12 derivações, *recall* de 99,93% e precisão de 99,13%.

Em Geovaents *et al* (2019) é desenvolvido um algoritmo para detectar e quantificar o complexo QRS fragmentado. A fragmentação QRS é um biomarcador que indica cicatrizes do miocárdio, detectável no ECG. A metodologia adotada para esta problemática possui quatro etapas: pré-processamento com filtros para remoção de ruídos e normalização do sinal, segmentação do complexo QRS utilizando a decomposição por modo variacional ou VMD (*Variational Mode Decomposition*), extração de características e classificação utilizando quatro algoritmos diferentes a saber: MVS, *K-nearest neighbors* (KNN), *Naive Bayesian* (NB) e *TreeBagger* (TB). Os algoritmos foram testados utilizando uma base de dados particular do Hospital Universitário de Leuven com 12 derivações, anotadas manualmente. O desempenho apresentou *recall* de 86% e especificidade de 89% com o algoritmo de MVS.

2.2.3 Trabalhos utilizando Aprendizado de Máquina

O uso do aprendizado de máquina aplicado a detecção do complexo QRS é inicialmente observado em tarefas não-supervisionadas com algoritmos de MVS (NAGAL e SHARMA, 2014). Apenas posteriormente é que se vê trabalhos publicados utilizando aprendizado supervisionado com redes neurais rasas (XIANG e LIN, 2018 e BELKADI *et al.*, 2021) e profundas. Nos detectores com redes profundas, há predominância das redes neurais convolucionais (*convolutional neural networks - CNNs*) (ZHONG *et al.*, 2018; SILVA, *et al*, 2018; BORDE *et al.*, 2019; WANG *et al.*, 2019; PARALIC, 2019; LIU *et al.*, 2020; JIA *et al.*, 2020; KOSCOVA *et al.*, 2021 e DURAJ *et al.*, 2022), seguido das arquiteturas de redes com memória de curto e longo prazo (*Long Short Term Memory Networks - LSTM*) (BORDE *et al.*, 2019), além de combinações entre elas (HABIB *et al.*, 2020 e BELKADI *et al.*, 2021).

Quanto ao uso de CNNs para a detecção do complexo QRS, conforme mostrado no Quadro 1, o modelo de WANG *et al.* (2019) apresenta *recall* de 99,92%, precisão de 99,98% e *F1-score* de 99,95%, superando os demais autores que fizeram uso da mesma abordagem. Neste trabalho o autor aplica etapas de pré-processamento como redução da taxa de amostragem, eliminação do ruído e normalização das amostras. A base de dados MIT-BIH com 46 amostras foi dividida

para as etapas de treinamento e teste, com anotações na posição dos picos R. Na etapa de pré-processamento a redução da amostragem, gerou amostras unidimensionais com 73 pontos. O valor de cada ponto desse vetor de amostras corresponde a amplitude do sinal de ECG filtrado e normalizado. A arquitetura do modelo é inspirada na arquitetura CNN ResNet, proposta em 2015, com a camada de convolução unidimensional. Nela o autor adiciona uma camada de ativação PReLU que aprende os parâmetros de inclinação usando retropropagação. O detector de WANG *et al.* (2019) consegue obter o pico R com alto desempenho, mas não é possível de estimar o intervalo de início e término do complexo QRS.

Outros autores (BORDE *et al.*, 2019 e DURAJ *et al.*, 2022) tiveram métricas com bons desempenhos e ambos faziam o processo de segmentação do sinal ECG, isto é, obtinham o intervalo do complexo QRS, da onda P e T. Em Borde *et al.* (2019) é testado 3 hipóteses de modelos (LSTM, 2-entradas 1D CNN e 2D CNN) e em Duraj *et al.* (2022) utiliza-se uma rede CNN do tipo U-net. A rede LSTM mostrou a impossibilidade de ser aplicado na vida real, mas a hipótese utilizando CNN com duas entradas unidimensionais de profundidades variadas apresentou as melhores métricas: *recall* de 99,36%, precisão de 98,99% e *F1-score* de 99,17%. O modelo testado por Duraj *et al.* (2022) na base de Lobachevsky University com 12 derivações do ECG obteve desempenho melhorado em relação à Borde *et al.* (2019) com *recall* de 99,45%, precisão de 99,27% e *F1-score* de 99,36%.

Dentre outras informações, no Quadro 1, para cada trabalho, são apresentados os métodos utilizados nas etapas de pré-processamento e de pós-processamento. A aplicação de filtros para a remoção de ruídos, remoção de picos imponentes, reamostragem do sinal e normalização são comuns no pré-processamento. Já no pós-processamento é empregado o limiar dos picos R-R e a altura média dos seguimentos QRS nas amostras detectadas após a rede neural.

Os trabalhos de Habib *et al.* (2020) e Belkadi *et al.* (2021) fazem uso de duas arquiteturas de aprendizado de máquina diferentes. No primeiro os autores utilizam uma rede CNN em série com uma rede LSTM, que realiza o pós-processamento. Os autores verificaram que uma CNN com 32 camadas de profundidade obteve a melhor taxa de *F1-score* em sete bases de dados diferentes. Particularmente, na base de dados INCART, essa métrica atingiu o valor de 91%. Em Belkadi *et al.* (2021) os autores usam uma arquitetura com *autoencoders* e *perceptron* multicamadas, sem pré

e pós-processamento. Para a base de dados INCART obteve-se um *recall* de 99,64%, precisão de 98,82 % e *F1-score* de 99,23%. A simplicidade da arquitetura utilizada torna-a adequada para ser utilizada em dispositivos embarcados.

Quadro 1 – Resumo da análise realizada nos trabalhos que tratam da detecção do complexo QRS em ECG.

Artigo	Base de Dados / Derivações	Abordagem	MÉTODOS			Resultados
			Pré-Processamento	Deteção do complexo QRS	Pós-Processamento	
DE OLIVEIRA, F. I. e CORTEZ, P. U.; 2004	MIT-BIH / Derivação Única	PDS	Não	Transformada de Hilbert sobre bases de Wavelet.	Não	Erro = 0,08%;
MEHTA, S. S.; LINGAYAT, N. S.; 2008	CSE / 1 e 12 derivações	PDS+ML	Remoção da linha de base utilizando FIR e das interferências da rede utilizando o filtro proposto por Furno e Tompkins (1983).	Critério de Entropia e MVS: Algoritmo K-Means.	Não	12 Derivações: Recall = 99,93% Precisão = 99,13% F1-score = 99,52%
HUANG, B.; WANG, Y.; 2009	INCART / 12 derivações	PDS	Remoção da Linha de base e do ruído de alta frequência utilizando filtro passa banda Butterworth.	PCA para reduzir a dimensão do filtro, Transformada Wavelet contínua e Entropia Wavelet combinada;	Não	Recall= 99,98% Precisão=99,99% F1-score = 99,98%
JALIL, B. et al.; 2010	MIT-BIH / Derivação Única	PDS	Remoção dos ruídos de interferência da rede e desvio da linha de base utilizando a transformada Wavelet Discreta.	Transformada Wavelet contínua e Limiar do módulo máximo.	Não	Variam com o tipo de amostra. Recall = 72,43% a 100% Precisão = 78,43% a 100%
LUTFULLAH, Z. et al.; 2011	INCART / 12 derivações	PDS	Remoção da linha de base utilizando técnicas baseadas em Wavelet.	Janela de Chebyshev e Limiar.	Não	Precisão = 99,4%
HU, Xiao et al.; 2014	MIT-BIH / 2 derivações	PDS	Não	Algoritmo de morfologia do triângulo modificado.	Não	Acurácia = 100%
NAGAL, D.; SHARMA, S.; 2014	CSE / 1 e 12 derivações	ML	Remoção da linha de base FIR e interferências da rede utilizando o filtro proposto por Furno e Tompkins (1983).	MVS: Algoritmo K-Means	Não	Acurácia= 99,89%
PANDIT, D. et al.; 2017	MIT-BIH, EDB, STDB, INCART e QTDB / 12 derivações	PDS	Suaviza o ruído e correção da linha de base utilizando a média da janela deslizante.	Máxima e Mínima Diferença (MMD)	Não	INCART Recall =98,64% Precisão = 99,31% F1-score = 98,97%
ZHONG, W. et al.; 2018	PCDB / 4 derivações	ML	Entropia da Amostra para identificação de ruído.	CNN	Não	Recall = 80,54%; Precisão = 75,33%; F1-score = 77,85%;
SILVA, P. et al.; 2018	MIT-BIH / Derivação única	ML	Não	CNN	Não	Recall = 99,35%; Precisão = 97,29%

Artigo	Base de Dados / Derivações	Abordagem	MÉTODOS			Resultados
			Pré-Processamento	Deteção do complexo QRS	Pós-Processamento	
						<i>F1-score = 98,31%</i>
LU, X. <i>et al.</i> ; 2018	MIT-BIH / Derivação única	PDS	Atenuação do ruído com filtro passa baixa.	Algoritmo de limiar adaptativo	Não	<i>Recall = 99,72%; Precisão = 99,69%; F1-score = 99,33%</i>
XIANG, Y.; LIN, Z.; 2018	MIT-BIH e INCART / 2 e 12 derivações	ML	Diferença entre as amostras adjacentes para acentuar o complexo QRS.	CNN+MLP	Não	<i>Recall = 99,86%; Precisão = 99,89%; F1-score = 99,87</i>
YU, R. <i>et al.</i> ; 2018	INCART / 12 derivações	ML	Redução da interferência da rede com o filtro de média e redução do efeito de desvio da linha de base aplicando filtro passa baixa <i>Butterworth</i> .	<i>Faster</i> R-CNN	Não	<i>Recall = 98,32% Precisão = 99,01% F1-score = 98,66%</i>
ZHAO, W. <i>et al.</i> ; 2018	INCART / 12 derivações	PDS	Remoção da linha de base e ruídos de alta frequência	Janela de fusão das derivações	Não	<i>Recall = 99,63% Precisão = 99,97% F1-score = 99,79%</i>
GOOVAERTS, G. <i>et al.</i> ; 2019	Base de dados do Hospital universitário de Leuven / 12 derivações	PDS+ML	Remoção do desvio da linha de base com filtro passa alta <i>Butterworth</i> ; Remoção do ruído de alta frequência utilizando filtro passa baixa e Normalização calculando <i>z-score</i> em cada derivação.	MVS e Decomposição do Modo Variacional (DMV)	Não	<i>Recall = 86% Especificidade = 89% Acurácia = 88%</i>
BORDE, A. <i>et al.</i> ; 2019	QTDB e MIT-BIH / 2 derivações	ML	Suavização com filtro de primeira derivada, Remoção de ruído e linha de base com filtro mediano FIR híbrido.	3 Hipóteses: 1º) LSTM 2º) 2-input 1-D CNN 3º) 2-D CNN	Filtro mediano das anotações	<i>Recall = 99,36% Precisão = 98,99% F1-score = 99,17%</i>
YUEN, B. <i>et al.</i> ; 2019	MIT-BIH e EDB / Derivação única	ML	Redução da Linha de base aplicando o filtro <i>Butterworth</i> passa alta e Normalização utilizando a média dos máximos locais.	CNN e LSTM	Não	<i>Recall = 98,73% Precisão = 97,35% F1-score = 98,03%</i>

Artigo	Base de Dados / Derivações	Abordagem	MÉTODOS			Resultados
			Pré-Processamento	Deteção do complexo QRS	Pós-Processamento	
WANG, X.; ZOU, Q. <i>et al.</i> ; 2019	MIT-BIH / 2 derivações	ML	Redução da amostragem em 180 Hz, eliminação do ruído de alta frequência e remoção do desvio da linha de base utilizando filtro <i>forward-backward</i> , normalização utilizando a diferença entre o limite superior e inferior das amostras.	CNN: ResNet	Não	<i>Recall</i> = 99,92% <i>Precisão</i> = 99,98% <i>F1-score</i> = 99,95%
PARALIC, M.; 2019	MIT-BIH / 2 derivações	ML	Não	CNN	Não	<i>Recall</i> = 83,7% <i>Precisão</i> = 85,7% <i>F1-score</i> = 83,9%
LIU, W. <i>et al.</i> ; 2020	CPSC2019, MIT-BIH NSTDB, QTDB e INCART / 1 e 12 derivações	ML.	Remoção de ruídos utilizando transformada Wavelet discreta.	CNN com convolução da oitava.	Limiar dos picos R	INCART <i>Recall</i> = 94,99% <i>Precisão</i> = 96,37% <i>F1-score</i> = 95,68%
HE, R. <i>et al.</i> ; 2020	MIT-BIH e CPSC2019 / 2 derivações	ML	Remoção da linha de base e remoção de ruídos utilizando transformada <i>Wavelet</i> .	CNN: U-net e LSTM	Limiar dos picos R	<i>Recall</i> = 99,56%; <i>Precisão</i> = 99,72% <i>F1-score</i> = 99,64%
CAI, W.; HU, D.; 2020	MIT-BIH, NSTDB e QTDB / 2 derivações	ML	Remoção de picos imponentes superior a 20mV, reamostragem do sinal utilizando transformação rápida de Fourie (para alcançar uma melhor generalização do modelo, a média dos valores do sinal é subtraída para cada gravação).	CNN e CRNN	Limiar de duração das amostras	<i>Recall</i> = 99,97% <i>Precisão</i> = 99,99% <i>F1-score</i> = 99,98%
JIA, M. <i>et al.</i> ; 2020	TELEDB, NSTDB e MIT-BIH / 1 e 2 derivações	ML	Reamostragem do sinal para 250 Hz e decomposição utilizando a decomposição de pacotes de <i>Wavelet</i> .	Time-Attention CNN	Eliminação das saliências utilizando um filtro mediano e binarização da amplitude por meio do limiar.	<i>Recall</i> = 99,85% <i>Precisão</i> = 99,92% <i>F1-score</i> = 99,88%

Artigo	Base de Dados / Derivações	Abordagem	MÉTODOS			Resultados
			Pré-Processamento	Deteção do complexo QRS	Pós-Processamento	
BELKADI, M. A. et al.; 2021	MIT-BIH, SVDB, INCART e EDB / 1, 2 e 12 derivações	ML	Não	Autoencoder e MLP	Não	<i>Recall</i> = 99,73% <i>Precisão</i> = 98,82% <i>F1-score</i> = 99,23%
FOTOOHINASAB, A. et al.; 2021	MIT-BIH / 2 derivações	ML	Não	Algoritmo de aprendizado de grafo	Não	<i>Recall</i> = 99,64% <i>Precisão</i> = 99,71% <i>F1-score</i> = 99,67%
HABIB, Ahsan. et al.; 2021	MIT-BIH, INCART, QTDB, EDB, STDB, TWADB, NSTDB e SVDB / 1, 2 12 derivações	ML	Reamostragem do sinal utilizando Nyquest, segmentação e sobreposição baseada no nível de decisão binária, normalização utilizando z-score.	CNN (gera a previsão) e RNN (corrige a previsão)	Filtragem do ruído de impulso e Limiar de distância R-R.	INCART: <i>F1-score</i> = 91%
KOSCOVA, Z. et al.; 2021	hospital FNUSA e hospital FNKV / 12 derivações	ML	Aumento de dados simulando batimentos cardíacos mais altos e Redução do sinal utilizando filtro <i>Butterworth</i> .	CNN: U-net	O ponto médio dos segmentos de amostras com probabilidades de QRS superior ao limiar de 0,955.	<i>Recall</i> = 96,49 % <i>Precisão</i> = 97,74 % <i>F1-score</i> = 97,12 %
DURAJ, K. et al.; 2022	LUDB / 12 derivações	ML	Não	CNN: U-net	Não	<i>Recall</i> = 95% <i>Precisão</i> = 95% <i>Especificidade</i> = 95%

LEGENDA:

PDS – Processamento Digital de Sinais;

ML – *Machine Learning* (Aprendizado de Máquina);

MIT-BIH – base de dados de arritmias do Instituto de Tecnologia de Massachusetts em parceria com o laboratório do hospital de Beth Israel em Boston;

CSE (*Common Standards for Electrocardiography*) base de dados com padrão comum de eletrocardiografia;

INCART (*Institute of Cardiological Technics*) – base de dados de arritmias do Instituto Técnico de Cardiologia de São Petersburgo;

EDB (*European ST-T Database*) – banco de dados europeu com registro das ondas ST e T disponibilizado pela Sociedade Europeia de Cardiologia;

STDB (MIT-BIH *ST Change Database*) – base de dados do MIT-BIH que exibem depressões transitórias dos segmentos de onda ST em ECG;

QTDB (QT-Database) – base de dados com seguimentos QT e QRS criado por meio das bases MIT-BIH e EDB;

PCDB (*PhysioNet Computing DataBase*) – base de dados do desafio da *PhysioNet* para localização do complexo QRS e intervalo QT em sinais de eletrocardiograma fetal;

CPSC2019 (*China Physiological Signal Challenge*) – base de dados para a detecção do complexo QRS e estimativa da taxa de batimento cardíaco em ECG de derivação única;

NSTDB (*Noise Stress Test DataBase*) – base de dados de teste de estresse de ruído produzidos nas gravações de ECG na base de dados do MIT-BIH.

TELDB (*Telehealth Database*) – base de dados

SVDB (*Supraventricular MIT-BIH Database*) – base de dados com análises da patologia de fibrilação supraventricular;

TWADB (*T-Wave Alternans Challenge Database*) – base de dados do desafio *PhysioNet* de alternância da onda T em ECG.

LUDB (*Lobachevsky University Electrocardiography Database*) – base de dados

FIR (*Finite Impulse Response*) - Filtro de Notch para resposta do impulso finito proposto por Van Alte e Schilder em 1985;

2.3 Discussão dos trabalhos

A detecção do complexo QRS em ECG possui dois grandes desafios. O primeiro é desenvolver uma abordagem que consiga detectar corretamente o intervalo do complexo QRS (verdadeiros positivos) e evitar regiões incorretas ou não QRS (falsos positivos), independentemente de qualquer patologia, derivação ou ruído que possa interferir na qualidade do sinal. O segundo, é desenvolver um algoritmo de detecção que seja leve, consiga ser executado nos dispositivos com baixo poder computacional e em tempo real.

Os primeiros trabalhos do Quadro 1, baseados em técnicas de PDS, representam um marco, por serem os primeiros a tentarem detectar o complexo QRS. É possível verificar que os métodos de PDS, na busca de uma melhoria do desempenho dos algoritmos, utilizaram abordagens variadas, como transformada *wavelet* e limiar do intervalo R-R.

Em geral, essas técnicas precisavam de pré-processamento, utilizando diversas técnicas baseada em filtros para tornar o sinal suave e linha de base corrigidos (HUANG e WANG, 2009; JALIL *et al.*, 2010; LUTFULLAH *et al.*, 2011; HU *et al.*, 2014), ou seja, isentos de ruídos de qualquer natureza. Em Jalil *et al.* (2010), os autores citam que muitos métodos foram desenvolvidos no passado para a detecção do complexo QRS, mas que as técnicas de remoção de ruído provocavam uma modificação do sinal original, o que produzia detecções falsas.

A etapa de pós-processamento não é identificada em nenhuma das técnicas de PDS. Na maioria, essas técnicas utilizam um limiar para detecção da onda R. No entanto, o uso de limiar dependia de um conjunto de regras empíricas ajustadas pelos autores (HUANG e WANG, 2009; JALIL *et al.*, 2010 e LUTFULLAH *et al.*, 2011). A aplicação dos algoritmos com PDS apresentou desempenhos baixos na detecção em intervalos de sinais que continham patologias ou em derivações diferentes.

A etapa de pré-processamento se manteve presente em todas as abordagens, e até mesmo nas de aprendizado de máquina, melhorando o desempenho de detecção dos algoritmos. No entanto, redundava em um custo computacional maior. Uma técnica de pré-processamento comumente utilizada pelos autores era fazer subamostragem do sinal, o que reduz o tempo de treinamento do modelo (WANG *et al.*, 2019).

As redes CNNs revolucionaram a forma como a detecção do complexo QRS era efetuada. A análise do Quadro 1 revela uma clara tendência das publicações a partir de 2018 de utilizarem redes CNNs. Xiang e Lin (2018) mostraram que o modelo proposto baseado em uma CNN rasa apresenta desempenho superior às pesquisas com abordagens de PDS. Isso se deve ao fato de essas redes detectarem o complexo QRS com exatidão, mas elas geram falsos positivos que são possíveis de eliminar fazendo pós-processamento do sinal (BORDE *et al.*, 2019).

No trabalho de (KOSCOVA *et al.*, 2021), para o treinamento de uma rede CNN tipo U-Net, os autores realizaram o treinamento da rede, apresentando na saída vetores binários, representando a probabilidade de encontrar um intervalo QS no mesmo. Um segmento identificado com o nível 1, representa um intervalo QS, enquanto um segmento com um nível 0 representa outros segmentos. A marcação do sinal QS foi feita a partir da marcação do pico R, marcando com o nível 1 as extensões do sinal, a partir desse ponto R, 0.05ms para a esquerda e 0.05ms para a direita. O intervalo de 0.05ms corresponde a 250 amostras do sinal. Essa marcação foi feita em três derivações precordiais, V1, V3 e V6. Uma etapa de pós-processamento elimina detecções incorretas.

A combinação de algoritmos de aprendizado de máquina não apresentou desempenho melhor de detecção do complexo QRS, quando comparado a detectores com apenas uma abordagem (BELKADI *et al.*, 2021). Além disso, na pesquisa de Habib *et al.* (2021), é provado que redes CNNs profundas não melhoram o desempenho de detecção, quando comparadas a redes mais rasas, ou seja, com menos camadas.

Os trabalhos mostram uma carência de base de dados com intervalos definidos de início e término do complexo QRS, o que dificultou obter com precisão a região de intervalo do complexo QRS. Obter a região do complexo QRS com exatidão é uma tarefa importante, pois várias patologias podem ser identificadas com base na altura e largura do intervalo.

3 REFERENCIAL TEÓRICO

O principal objetivo desse trabalho é utilizar Redes Neurais Convolucionais para detectar o complexo QRS em sinais ECG com 12 derivações. Para facilitar o entendimento do que foi realizado, apresentamos a seguir, com brevidade, conceitos fundamentais a respeito da anatomia e fluxo sanguíneo no coração e dos sinais elétricos responsáveis pelo batimento cardíaco. São apresentados também conceitos relacionados as redes neurais convolucionais, utilizadas nesse trabalho para a detecção do complexo QRS.

3.1 O Eletrocardiograma

O eletrocardiograma (ECG ou EKG) é um tipo especial de gráfico que representa a atividade elétrica do coração. Especificamente, o ECG fornece um mapa de tensão versus o tempo dos batimentos cardíacos. O ECG é uma parte importante do diagnóstico clínico e manejo de pacientes internados e ambulatoriais, pois pode fornecer informações críticas (GOLDBERGER, 2018).

O coração humano é um órgão muscular formado por quatro câmaras ocas que o dividem em duas porções, sendo a superior composta pelos átrios esquerdo e direito e a inferior pelos ventrículos esquerdo e direito. A função central do coração é contrair ritmicamente e bombear sangue para os pulmões (circulação pulmonar) para oxigenação e depois bombear esse sangue enriquecido com oxigênio para a circulação geral (sistêmica).

A quantidade de sangue bombeada deve ser compatível com as diferentes necessidades metabólicas do corpo. Nosso músculo cardíaco e outros tecidos precisam de mais oxigênio e nutrientes quando estamos ativos do que quando estamos em repouso. Uma parte importante desses ajustes autorregulatórios é feita por meio de alterações na frequência cardíaca, que estão amplamente sob o controle do sistema nervoso autônomo (involuntário).

O sinal de contração cardíaca é a propagação de correntes elétricas sincronizadas através do miocárdio. Essas correntes são produzidas pelas células marca-passo e por tecido condutor especializado dentro do coração, bem como pelo próprio músculo cardíaco em funcionamento.

As células marca-passo são como pequenos osciladores que geram automaticamente estimulação elétrica de maneira repetitiva. Outras células cardíacas, incluindo tecido condutor especializado e miocárdio funcional, atuam como cabos que transmitem esses sinais elétricos.

As correntes elétricas são produzidas pela disseminação sequencial e organizada de estímulos através dos átrios e ventrículos, seguidos de seu retorno ao estado de repouso. Além disso, cada fase da atividade elétrica cardíaca produz uma onda ou deflexão específica, conforme mostra a Figura 2. As cinco formas de onda básicas de ECG, são:

- Onda P – representa a atividade elétrica dos átrios durante a contração, é observada antes do complexo QRS sendo pequena em amplitude. Tem duração de cerca de 0,12 a 0,20 segundos.
- Complexo QRS – representa a atividade elétrica dos ventrículos durante a contração, tem uma duração de 0,06 a 0,10 segundos. É observado logo após a onda P.
- Segmento ST – representa a atividade elétrica dos ventrículos entre a onda S do complexo QRS e a onda T. Pode indicar se há problemas com a atividade elétrica dos ventrículos.
- Onda T – representa a recuperação dos ventrículos após a contração. Tem duração de cerca de 0,20 a 0,40 segundos. É observada logo após o complexo QRS.
- Onda U – é observada em algumas condições patológicas, como bloqueios de ramo direito ou lesão miocárdica.

3.1.1 Derivações do ECG

A cada ciclo cardíaco, as estruturas envolvidas na condução irão gerar potenciais de ação que carregam formas de onda. A contribuição de cada potencial de ação ao longo do tempo gera potenciais elétricos em várias partes do corpo. Como os músculos ventriculares e átrios compõem a maior parte do tecido cardíaco, suas atividades se destacam.

Durante a passagem do estímulo elétrico pelo coração, uma fração da corrente elétrica flui pelo corpo criando potenciais elétricos em diferentes pontos, que podem ser registrados por eletrodos. Para a obtenção do eletrocardiograma, foram exploradas localizações padronizadas no corpo humano, que formaram as derivações eletrocardiográficas. As derivações são divididas em duas, derivações periféricas e derivações unipolares precordiais. A primeira utiliza os membros superiores e inferiores como pontos de aquisição dos potenciais, gerando as derivações *I*, *II*, *III*, *aV_R*, *aV_L* e *aV_F*. Já a segunda mede os potenciais através de eletrodos espalhados no tórax do indivíduo provendo as derivações *V₁*, *V₂*, *V₃*, *V₄*, *V₅* e *V₆*.

3.1.2 Complexo QRS

O complexo QRS representa o grupo de ondas que traduzem a atividade ventricular correspondendo a despolarização da musculatura cardíaca. Ele é composto de três ondas: a onda Q, a onda R e a onda S, conforme a Figura 2. A onda Q indica a despolarização do septo interventricular, a onda R representa a despolarização da parede anterior do ventrículo esquerdo e a onda S representa a despolarização da parede posterior do ventrículo esquerdo.

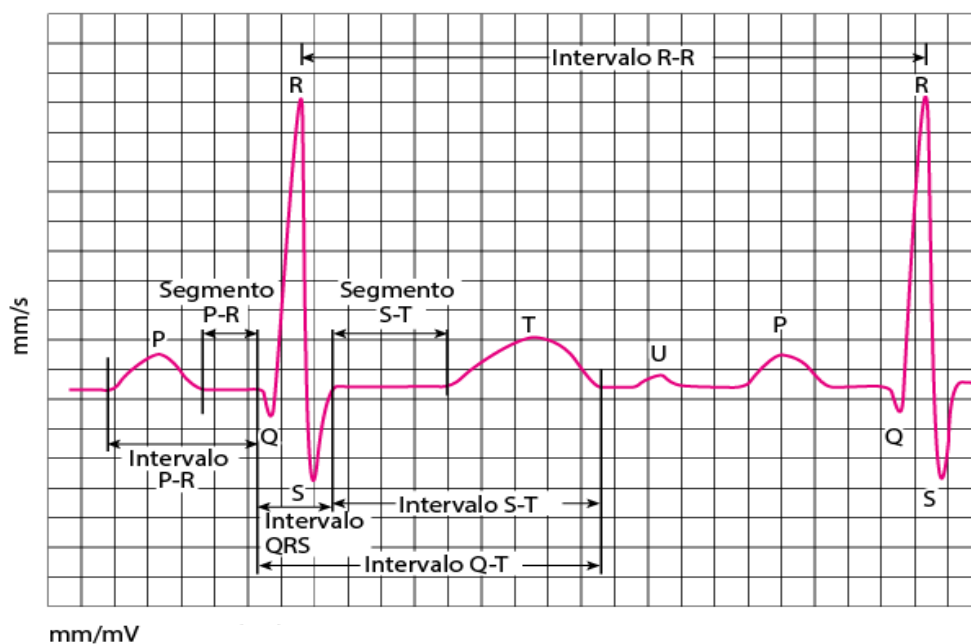


Figura 2: Representação do intervalo QRS em um sinal ECG.
Fonte: Adaptado de Goldberger (2018).

De acordo com Goldberger (2018) há regras básicas para nomear os componentes do complexo QRS em qualquer derivação do eletrocardiograma. A nomenclatura do complexo QRS é baseada nas deflexões negativas e positivas no sinal, a onda Q é a deflexão inicial negativa, a onda R é a primeira deflexão positiva e a onda S é a deflexão negativa após a onda R. Nem todo complexo QRS contém todas as três ondas, o que pode levar a confusão na nomenclatura.

A Figura 3 mostra os principais complexos QRS possíveis e a nomenclatura das respectivas ondas. As letras maiúsculas são usadas para designar ondas de amplitude relativamente grande e as minúsculas para ondas relativamente pequenas.

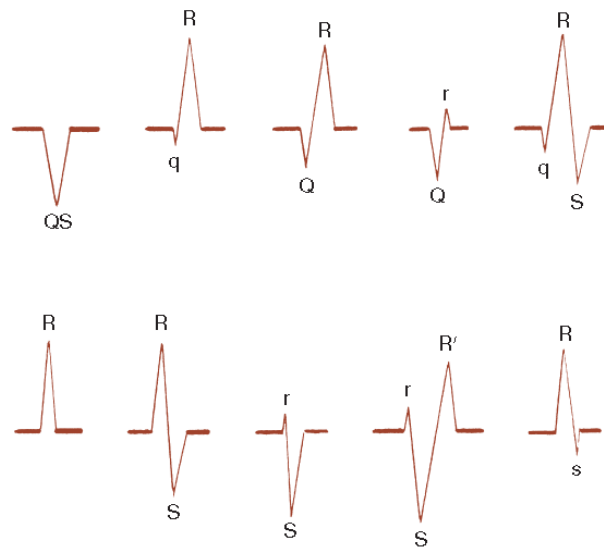


Figura 3: Diferentes apresentações do complexo QRS.
Fonte: Goldberger (2018).

3.2 Redes Neurais Artificiais (RNA)

As Redes Neurais Artificiais (RNA) são modelos computacionais inspirados no sistema nervoso de seres vivos em especial na forma como o cérebro desempenha uma tarefa. Estas redes possuem unidades de processamento chamadas de neurônios e que estão interligadas entre si. A força das interconexões entre os neurônios gera pesos sinápticos proporcionando a capacidade de uma rede aprender a desempenhar uma tarefa.

O estudo das RNA inicia com a proposta de McCulloch e Pitts (1943) que simulam modelos de neurônios cerebrais na forma de algoritmos binários criando a base para a modelagem de sistemas neurais. Em seguida, várias propostas tornaram

o uso das RNA aplicáveis, tais como: o conceito de aprendizagem por reforço utilizando modelos matemáticos, a primeira rede neural de camada única chamada *Perceptron*, o treinamento das redes *perceptron* de multicamadas utilizando *backpropagation* e o treinamento das RNA aplicado à imagens usando o método de convolução com unidades gráficas de processamento ou GPU. A Figura 4 exemplifica uma RNA com várias entradas e várias saídas.

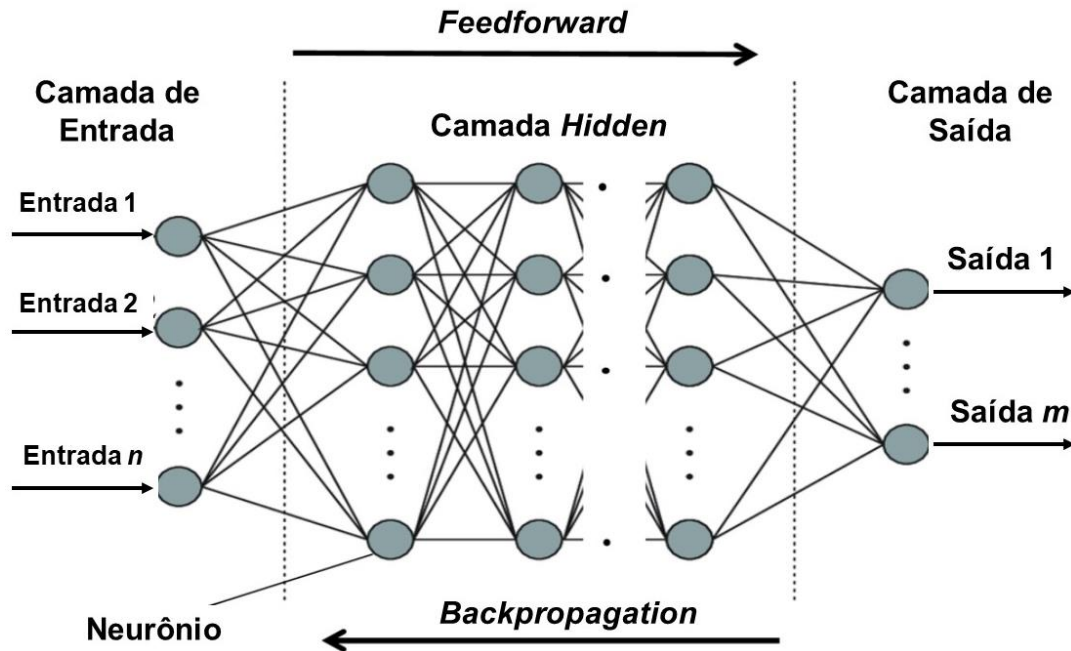


Figura 4: Rede Neural Artificial totalmente conectada (densa).
Fonte: Autor

Um *perceptron* de multicamada representa um modelo adaptável $y(\cdot, w)$ capaz de mapear entradas com D -dimensões para saída(s) com C -dimensões:

$$y(\cdot, w): \mathbb{R}^D \rightarrow \mathbb{R}^C, x \mapsto y(x, w) = \begin{pmatrix} y_1(x, w) \\ \vdots \\ y_C(x, w) \end{pmatrix} \quad (1)$$

Em geral, um *perceptron* de $(L + 1)$ camadas consiste em $(L + 1)$ camadas, cada camada l processando combinações lineares da camada anterior $(l - 1)$ (ou a entrada). Na entrada $x \in \mathbb{R}^D$, um neurônio artificial consiste em uma combinação linear de entradas do tipo x_1, x_2, \dots, x_n na forma:

$$z = \sum_{i=1}^n w_i x_i + w_o \quad (2)$$

Onde w_i representa os pesos ajustáveis e w_o é chamado de viés (*bias*). A saída z deste neurônio é dado por uma função de ativação diferenciável não linear $f(\cdot)$ sobre z :

$$y = f(z) \quad (3)$$

A saída calculada de y da i –ésima unidade computacional de uma camada l da rede neural $(L + 1)$ camadas é representa por:

$$y_i^{(l)} = f(z_i^{(l)}) \text{ com } z_i^{(l)} = \sum_{j=1}^D w_{i,j}^{(l)} x_j + w_{i,0}^{(l)} \quad (4)$$

3.2.1 Funções de Ativações

Função de ativação é um dos blocos de construção da Rede Neural. Elas aprendem os recursos abstratos por meio de transformações não lineares e assim determinam a saída de um modelo, sua precisão e eficiência computacional. Em alguns casos, as funções de ativação têm um efeito importante na capacidade de convergência do modelo e na velocidade de convergência (Ekman, 2021). Há uma variedade de funções de ativação, mas comumente utiliza-se a função tangente hiperbólica:

$$\tanh z = \frac{e^z - e^{-z}}{e^z + e^{-z}} \quad (5)$$

Outra função muito utilizada é a função sigmoide:

$$\sigma(z) = \frac{1}{1 + e^{-z}} \quad (6)$$

3.2.2 Treinamento da rede

Durante o processo de treinamento da rede neural, para cada exemplo de treinamento no conjunto de $n = 1, 2, \dots, N$ elementos, apresenta-se o vetor de característica x_n para a rede neural e compara-se a saída prevista por ela $y(x_n; w)$ com a correspondente saída esperada t_n . Logo, a função de perda mede o quanto a saída prevista da rede é diferente da saída esperada (CHARU, 2018).

Para problemas de classificação é utilizada a função de perda de entropia cruzada e para problemas de regressão usa-se a função de perda quadrática. A função de erro muito utilizada é a soma dos quadrados do erro, $E(w)$:

$$E(w) = \sum_{n=1}^N E_n(w) = \sum_{n=1}^N \sum_{k=1}^C (y_k(x_n; w) - t_{nk})^2 \quad (7)$$

O processo de treinamento da rede busca encontrar um conjunto de pesos de modo que minimize a função indicada na equação (7). Os algoritmos utilizados para a minimização da função de erro baseiam-se no uso de derivadas do vetor gradiente $\nabla E(w)$. O algoritmo de retropropagação (*backpropagation*) é uma técnica eficiente para calcular o gradiente em redes neurais, este algoritmo ocorre da seguinte forma:

1. Para um vetor de entrada x_n , aplica-se o caminho direto (*Feedforward*) para calcular as ativações e saídas para todas as camadas da rede;
2. Calcular as funções de erro δ_k a partir das unidades de saída, utilizando:

$$\delta_k = y_k - t_k \quad (8)$$

3. Calcular as funções de erro δ_j de cada unidade escondida no sentido contrário usando o algoritmo de retropropagação:

$$\delta_j^{(l)} = h'(z_j^{(l)}) \sum_k w_{kj}^{(l+1)} \delta_k^{(l+1)} \quad (9)$$

4. Estimar as derivadas necessárias, utilizando:

$$\frac{\partial E_n}{\partial w_{kj}^{(l)}} = \delta_j^{(l)} z_i^{(l-1)} \quad (10)$$

5. Alterar os pesos com base nos gradientes estimados por $-a \cdot \nabla E_n(w)$, onde a representa a taxa de aprendizagem;
6. Voltar para o passo 1. e repetir até um número de iterações ou um mínimo desejado.

3.3 Redes Neurais Convolucionais

As Redes Neurais Convolucionais (CNN) é um tipo de rede neural profunda que vem sendo cada vez mais utilizada. Ela é composta por camadas que realizam operações lineares e não lineares. A proposta inicial dessa rede data de 1998 por LeCun *et al.* (1998), mas foi a partir da última década que ela passou a ser mais

utilizada garantindo melhor desempenho de classificação. Desde então a CNN vem alcançando o estado da arte em várias aplicações relacionadas principalmente ao processamento de imagens.

Geralmente, as CNNs classificam as imagens apresentadas a elas. Elas são formadas por neurônios com conexões contendo pesos e polarizações, que são ajustadas durante as interações, de forma semelhante às RNAs discutidas anteriormente.

Ao contrário das RNAs, os neurônios das Redes Neurais Convolucionais são conectados apenas a alguns neurônios na camada anterior. Portanto, os neurônios da CNN se especializam em apenas uma pequena área da imagem. Devido a um número reduzido de conexões, eles podem lidar com um número muito alto de entradas (CHARU, 2018).

Cada camada de redes neurais profundas é composta por várias camadas de processamento não linear, onde a saída de uma camada alimenta a entrada da próxima. Como resultado de várias camadas de processamento, uma rede neural pode atingir níveis mais altos de abstração (como identificação de objetos) com recursos de baixo nível (como pixels). Como resultado, esse tipo de classificador elimina os estágios de seleção e extração de recursos.

Figura 5 fornece uma visão ampla da aparência de uma CNN, as principais camadas. As camadas estão agrupadas em dois blocos de construção básicos, o bloco de extração de características e o bloco de classificação.

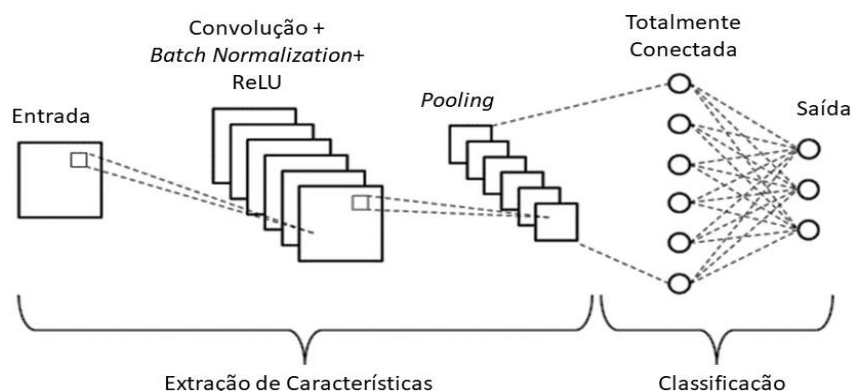


Figura 5: Arquitetura de uma CNN contendo as principais camadas de processamento.

Fonte: adaptado de Géron (2022).

O primeiro bloco consiste das camadas de Convolução, *Batch Normalization*, Unidade Linear Retificadora (*ReLU*) e a camada de *Pooling*. Essas camadas formam o componente essencial da extração de características. O segundo bloco consiste nas camadas totalmente conectada. Esta camada executa a tarefa de classificação com base na entrada do bloco convolucional.

3.3.1 Camada Convolutiva

A camada convolutiva caracteriza-se pela primeira camada em uma CNN representando o componente-chave da arquitetura. Realiza a operação matemática linear entre duas funções chamada convolução.

Ela é realizada deslizando um elemento estruturante (também chamado por máscara, filtro ou *kernel*) sobre cada pixel da imagem, dessa forma gera-se como resultado um mapa de características (GONZALES, 2018). A definição matemática da operação de convolução é representado pela equação (11).

$$g(x, y) = I(x, y) * f(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b I(s, t) * f(x - s, y - t) \quad (11)$$

Nessa equação $I(x, y)$ representa a imagem com uma ou mais matrizes 2D, onde conexões em regiões pequenas e localizadas são feitas através do elemento estruturante $f(x, y)$.

A exemplificação de uma operação de convolução é apresentada na Figura 6. Nela há uma matriz que corresponde a uma imagem em tons de cinza de dimensão 5 x 5, e um elemento estruturante com dimensões 2 x 2. A resposta desse processo de convolução é um mapa de características de dimensão 4 x 4, este valor é obtido por meio da somatória das multiplicações dos valores do elemento estruturante pelos respectivos valores da imagem.

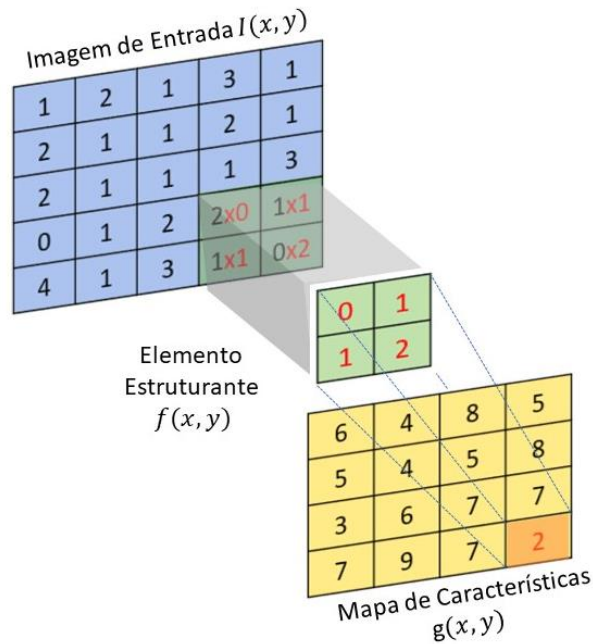


Figura 6: Ilustração da operação de convolução linear.
 Fonte: Adaptado de Dalvi (2018).

O *kernel* é formado por pesos que são inicializados aleatoriamente e atualizados pelo método de propagação reversa (*backpropagation*), a cada etapa do treinamento, quando submetido a um novo padrão na entrada da rede (GONZALES, 2018).

3.3.2 Camada *Batch Normalization*

A camada de *Batch Normalization* (BN) ou normalização em lote vem depois de uma camada de convolução e antes de uma camada de ReLU, seu objetivo é acelerar o treinamento e reduzir a sensibilidade em relação a inicialização de pesos e polarizações da rede. Esta camada é aplicada de modo a evitar problemas de gradiente de fuga/explosão durante o treinamento.

O efeito de fuga/explosão ocorre durante a etapa de *backpropagation* quando o gradiente descendente torna inalterado os pesos nas camadas inferiores e o treinamento nunca converge para uma solução satisfatória (gradiente de fuga), ou quando o gradiente aumenta, tornando as atualizações extremamente grande divergindo do algoritmo (explosão do gradiente).

A BN realiza isso por meio de uma etapa de normalização que corrige as médias e variações das entradas da camada, assim, permite-se o uso de taxas de

aprendizado muito mais altas sem o risco de divergência. Com o objetivo de centralizar em zero e normalizar as entradas, o algoritmo da camada de BN precisa estimar a média e o desvio padrão de cada entrada em relação ao *mini-batch* (pequenos lotes). O algoritmo de *Bach Normalization* segue o passo a passo conforme as equações de 12 a 15:

- 1) μ_B é o vetor de médias da entrada calculado em parcelas de *mini-batch* e m_B é o número de instâncias no *mini-batch*:

$$\mu_B = \frac{1}{m_B} \sum_{i=1}^{m_B} x^{(i)} \quad (12)$$

- 2) σ_B representa o vetor das variâncias de entrada calculado em parcelas de *mini-batch*:

$$\sigma_B^2 = \frac{1}{m_B} \sum_{i=1}^{m_B} (x^{(i)} - \mu_B)^2 \quad (13)$$

- 3) $\hat{x}^{(i)}$ apresenta o vetor de entradas normalizadas e centralizadas em zero para a instância i , a divisão por zero é evitada utilizando um número ínfimo ε que normalmente é 10^{-5} :

$$\hat{x}^{(i)} = \frac{x^{(i)} - \mu_B}{\sqrt{\sigma_B^2 + \varepsilon}} \quad (14)$$

- 4) $z^{(i)}$ é a saída da operação de BN, obtida por meio da multiplicação (\otimes) elemento a elemento do vetor de parâmetros do escalonamento de saída γ para a camada pelo resultado do passo (3), somado ao vetor de parâmetros de deslocamento de saída β da camada:

$$z^{(i)} = \gamma \otimes \hat{x}^{(i)} + \beta \quad (15)$$

3.3.3 Camada de Unidades Lineares Retificadoras (ReLU)

A *Rectified Linear Unit* (ReLU) ou Unidade Linear Retificadora é uma função de ativação amplamente utilizada em redes neurais profundas devido à sua simplicidade e eficácia. No Tópico 3.2.1 apresentou-se as funções de ativações, as quais são utilizadas em uma rede neural para controlar a saída de cada neurônio de acordo com a entrada permitindo ao modelo aprender relações não lineares.

A utilização da ReLU ajuda a resolver o problema de "*vanishing gradient*" em redes neurais profundas. Ela impede que os pesos das camadas anteriores sejam atualizados com valores muito pequenos ou negativos, o que pode ocorrer com outras funções de ativação, como a sigmoide. Isso ajuda a garantir que a rede possa aprender com eficácia e atualizar seus pesos de forma apropriada.

A ReLU é definida como na equação (16), onde z é a entrada da unidade. Isso significa que a saída da unidade é igual ao valor de entrada quando este é positivo e zero quando é negativo.

$$g(z) = \max\{0, z\} \quad (16)$$

A ReLU é definida como na equação (16), onde z é a entrada da unidade. Isso significa que a saída da unidade é igual ao valor de entrada quando este é positivo e zero quando é negativo.

3.3.4 Camada de *Pooling* (ou sub-amostragem)

A camada de *pooling* é uma técnica utilizada em redes neurais convolucionais para reduzir a dimensionalidade dos mapas de características e aumentar a robustez a variações de posição e escala. É também conhecida como sub-amostragem ou sub-agrupamento (*pooling*).

De acordo com (GÉRON, 2022), a camada de *pooling* é responsável por selecionar as características mais importantes de cada janela de entrada, permitindo que a rede aprenda características mais abstratas e robustas. Esse processo também ajuda a evitar *overfitting*, pois reduz o número de parâmetros da rede.

Existem vários tipos de *pooling*, sendo o mais comum o máximo (*max-pooling*) e média (*average-pooling*). A Figura 7 apresenta exemplos dos procedimentos mais comuns de sub-agrupamentos. O *max-pooling* seleciona o valor máximo de cada janela de entrada, enquanto o *average-pooling* seleciona a média dos valores. Segundo (ZHOU, 2021), o *max-pooling* é mais comumente utilizado devido ao seu bom desempenho em tarefas de reconhecimento de objetos.

Além disso, (EKMAN, 2021) aponta que as camadas de *pooling* também podem ser usadas para aumentar a invariância à translação, tornando a rede capaz de detectar objetos em diferentes posições na imagem de entrada.

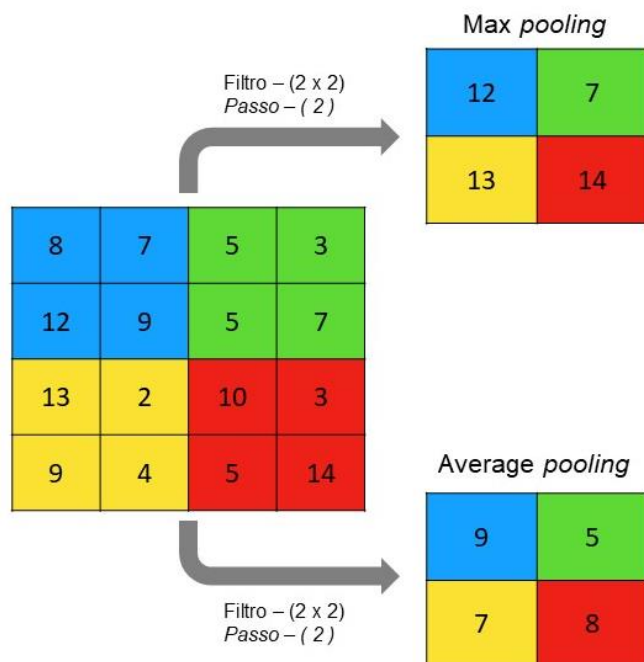


Figura 7: Exemplo da operação matemática na camada de *pooling*.
Fonte: Adaptado de ZHANG *et al.* (2021)

Em resumo, a camada de *pooling* é uma técnica fundamental em redes neurais convolucionais, pois ajuda a reduzir a dimensionalidade dos mapas de características, aumenta a robustez a variações de posição e escala e ajuda a evitar *overfitting*. É importante citar que a operação de *pooling* não altera a profundidade da rede neural.

3.3.5 Camada de Convolução Transposta (ou sobre-amostragem)

A camada de Convolução Transposta, também conhecida como deconvolução ou sobre-amostragem, é uma camada de rede neural utilizada para aumentar a resolução de mapas de recursos bidimensionais. Ela funciona como uma camada de Convolução inversa, aumentando a escala dos dados de entrada e expandindo seus mapas de características.

A operação de convolução padrão reduz a amostra da entrada aplicando filtros convolucionais deslizantes. A operação de convolução é expressa como na equação (17) para a matriz de convolução C e o vetor de viés B que pode ser derivado dos pesos e vieses da camada.

A operação de Convolução Transposta é utilizada para aumentar a resolução da entrada, ao contrário da Convolução tradicional que é utilizada para diminuir a resolução. Isso é alcançado aplicando filtros convolucionais deslizantes à entrada e

preenchendo com zeros as bordas da entrada com preenchimento que tem o tamanho da borda do filtro correspondente menos 1. Dessa forma, a Convolução Transposta adiciona detalhes à entrada original, aumentando a sua resolução.

$$Y = C X + B \tag{17}$$

A operação de convolução transposta é equivalente a equação (18) onde C e B denotam a matriz de convolução e o vetor de viés para a convolução padrão derivada dos pesos e vieses da camada, respectivamente.

$$Y = C^T X + B \tag{18}$$

A Figura 8 apresenta um exemplo que demonstra a operação de convolução transposta com um elemento estruturante 2 x 2 aplicados à uma entrada 2 x 2.

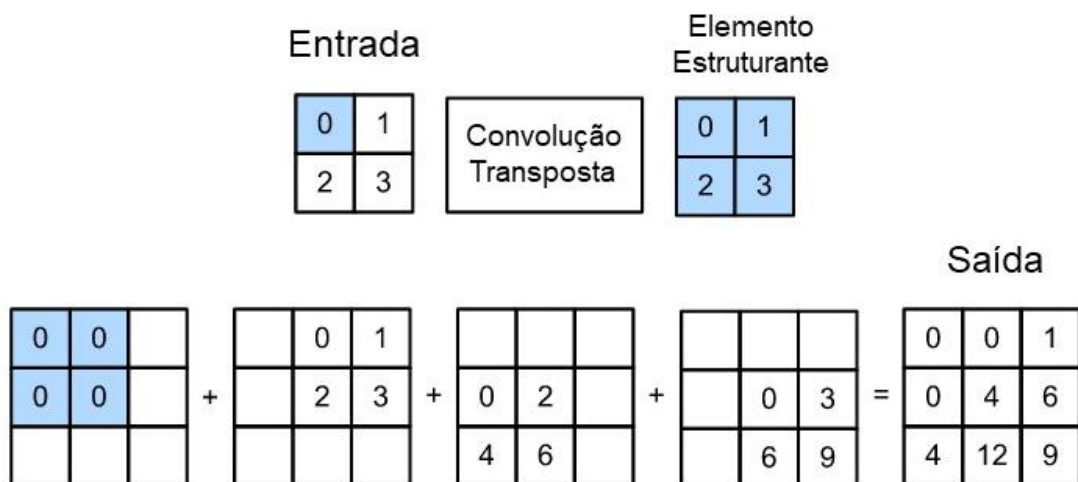


Figura 8: Modelo de Operação da convolução transposta.
Fonte: Adaptado de ZHANG *et al.* (2021)

3.3.6 Camada de Classificação

A camada de classificação, que é usada ao final da arquitetura de uma CNN, tem como objetivo estimar uma saída para a rede. Durante o treinamento, é calculada a diferença entre a saída estimada e o valor desejado para que os parâmetros de aprendizado possam ser ajustados de maneira a minimizar essa diferença.

A função *Softmax* é geralmente utilizada neste contexto, pois é a mais adequada quando a tarefa de classificação envolve várias classes. Esta função tem a vantagem de fornecer saídas que somam 1 e são positivas, permitindo interpretar a

saída da rede como uma distribuição de probabilidade discreta da entrada pertencer a cada classe. Esta probabilidade é representada pela equação (19), onde z é o valor da ativação da entrada da camada *Softmax* definido na equação (4).

$$p_i = \frac{e^{z_i}}{\sum_j^n e^{z_j}} = y_i \quad (19)$$

A classe estimada \hat{i} será aquela de maior probabilidade

$$\hat{i} = \arg \max (p_i) \quad (20)$$

Para as redes neurais convolucionais, após a aplicação da função *Softmax*, é feito o teste da confiabilidade do modelo utilizando a função de entropia cruzada (*cross-entropy*) para o cálculo da perda (*loss*), com o objetivo de maximizar o desempenho da rede neural. O cálculo da perda é definido como:

$$H(p, q) = -\frac{1}{N} \sum_{n=1}^N \sum_{i=1}^K w_i t_{ni} \log y_{ni} \quad (21)$$

Onde N é o número de amostras, K o número de classes, w_i é o peso para a classe i , t_{ni} é o indicador de que a n -ésima amostra pertence à i -ésima classe, y_{ni} é a saída da amostra n para a classe i , que neste caso, é o valor da função softmax. Portanto, y_{ni} é a probabilidade de que a rede associe a n -ésima entrada à classe i .

3.3.7 Camada *Dropout*

O *Dropout* é uma técnica de regularização eficaz para redes neurais. O procedimento é feito removendo aleatoriamente um subconjunto de neurônios da rede durante o treinamento. O subconjunto de neurônios removidos varia ao longo de cada época de treinamento.

Esta técnica força as unidades da rede neural a trabalhar com conjuntos aleatórios de outras unidades, evitando assim que subconjuntos de unidades se adaptem para resolver casos específicos e demonstrou reduzir o *overfitting*.

A Figura 9 demonstra o efeito nos neurônios com e sem a utilização de *Dropout*, observa-se que ele é adicionado como uma camada após a camada onde se deseja aplicá-lo. A camada *dropout* bloqueia as conexões de um subconjunto de neurônios da camada anterior, o que tem o mesmo efeito como se o neurônio não estivesse lá.

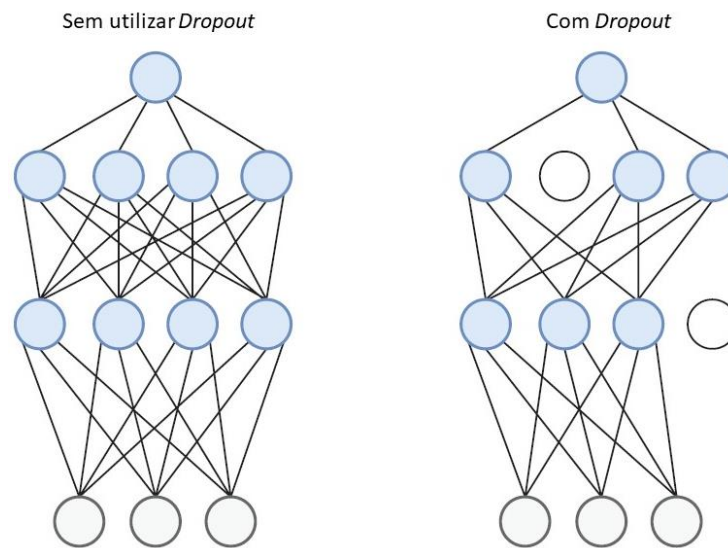


Figura 9: Exemplo do efeito nas conexões dos neurônios com e sem o uso de *dropout*.
Fonte: (EKMAN, 2021)

3.4 Processo de Treinamento de uma Rede Convolutiva

O treinamento de redes neurais convolutivas é baseado no algoritmo de *Backpropagation*. Nele, quando uma entrada é apresentada à rede, esta faz uma predição da saída e caso haja diferença entre o predito e o desejado, os parâmetros internos da rede sofrem modificações, na tentativa de minimizar o erro.

O *Backpropagation* é um dos algoritmos de otimização mais utilizados para treinar redes neurais, e se baseia na ideia de calcular o gradiente da função de perda (ou erro) e usá-lo para atualizar os pesos das conexões na rede, a fim de minimizar a perda. O processo de atualização dos pesos é feito com base na informação fornecida pela passagem da entrada através da rede na direção *forward*, e ajustando os pesos na direção *backward*. Conforme descrito no item 3.2.2 o treinamento de uma rede neural convolutiva segue os passos indicados de acordo com o método de otimização definido nos parâmetros de treinamento do modelo.

3.4.1 Método de Otimização

Os métodos de otimização são usados para treinar redes neurais. Eles ajudam a minimizar a função de custo ou erro da rede, logo são melhoramentos ou variações do algoritmo de *backpropagation*. Cada conjunto de dados pode ter melhores resultados com um método de otimização específico, portanto não há um método que

seja claramente o melhor, mas sim aqueles que se adaptam melhor ao problema. Há três métodos de otimização comumente usados, que são SGDM (*Stochastic Gradient Descent With Momentum*), RMSProp (*Root Mean Square Propagation*) e ADAM (*Adaptive Moment Estimation*). É importante entender como o algoritmo de treinamento afeta os resultados da rede neural e escolher o método correto pode significar uma grande diferença na velocidade e qualidade dos resultados obtidos.

O método ADAM, é uma extensão do SGD. Ele foi apresentado por Diederick Kingma (2014) e é um dos algoritmos mais utilizados para o treinamento de redes neurais profundas, os autores (KINGMA; BA, 2014) demonstraram o ADAM sendo superior aos métodos de otimização mais utilizados em redes neurais convolucionais. A Figura 10 mostra o resultado da comparação com outros métodos de otimização obtido pelos autores na base de dados MNIST (base de dados de números escritos a mão), na tarefa de classificação.

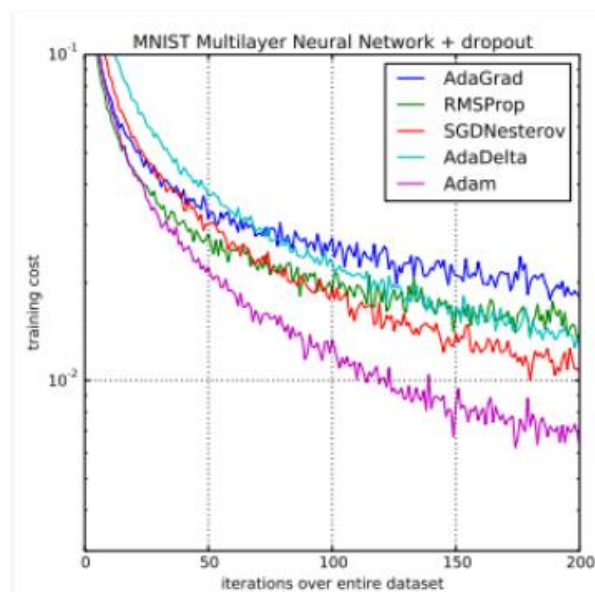


Figura 10: Comparação do ADAM com outros métodos de otimização. Fonte: (KINGMA; BA, 2014)

3.4.2 Segmentação Semântica

As redes neurais convolucionais (CNNs) são amplamente utilizadas para tarefas de classificação local, também conhecidas como segmentação semântica. A segmentação semântica é uma tarefa que consiste em atribuir uma classe a cada pixel

de uma imagem. Isso pode ser usado para dividir uma imagem em diferentes regiões, como objetos, fundo, etc. A Figura 11 exemplifica esse procedimento.

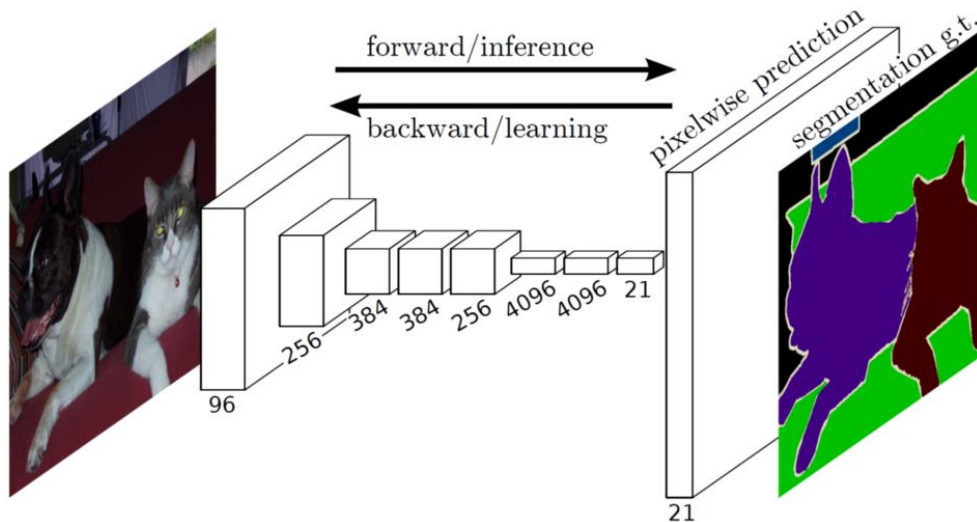


Figura 11: Exemplo de uma CNN para a segmentação semântica.
Fonte: (LONG, SHELHAMER e DARRELL, 2015).

Um dos processos utilizados na segmentação semântica é a convolução transposta, detalhada no item 3.3.5. A convolução transposta é uma camada da rede neural de aprendizado profundo que visa ampliar a área de entrada, aumentando a resolução espacial das características extraídas anteriormente. De acordo com os autores Long, Shelhamer e Darrell (2015), "a convolução transposta é um dos blocos-chave da arquitetura de uma rede neural para a segmentação semântica".

A convolução transposta ajuda a preservar as características da imagem original, permitindo que a rede identifique diferentes objetos ou áreas dentro da imagem. Isso é possível porque a convolução transposta inverte o processo de convolução, ampliando as características importantes para o processo de segmentação.

4 MATERIAIS E MÉTODOS

O presente trabalho propõe um método para detecção automática do complexo QRS em sinal de eletrocardiograma (ECG) com 12 derivações. No Capítulo 3 foram apresentados os principais conceitos teóricos que subsidiam a implementação desta pesquisa. Neste capítulo serão expostos os materiais necessários para esta implementação, os passos da metodologia e as arquiteturas de redes CNNs propostas. Para alcançar os objetivos geral e específicos desta pesquisa, a metodologia adotada utilizou as etapas apresentadas no diagrama em blocos da Figura 12.

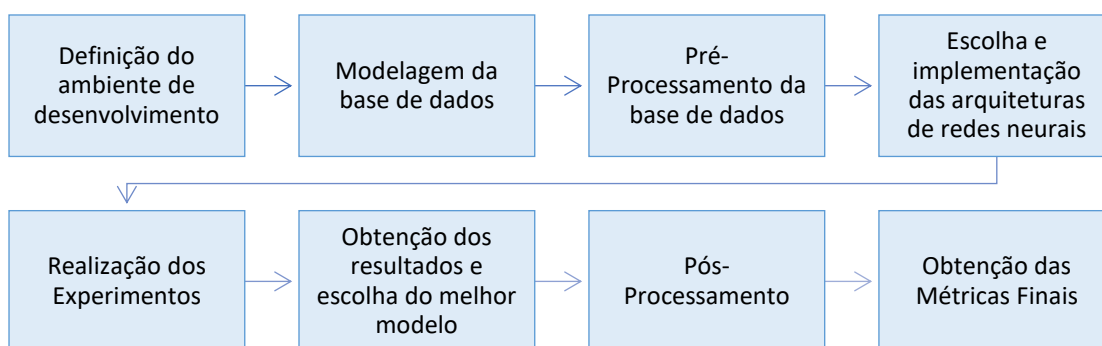


Figura 12: Diagrama em blocos apresentando as etapas da metodologia utilizada nesse trabalho para detecção do complexo QRS.

Na Figura 12, a primeira etapa corresponde a definição do ambiente de desenvolvimento da pesquisa, onde são descritas as ferramentas de hardware e software utilizadas para implementação do treinamento, validação e teste dos classificadores. A segunda etapa corresponde a modelagem da base de dados com os sinais das 12 derivações do ECG. Nessa etapa é detalhado como foram construídos os conjuntos de treinamento, validação e teste.

Em seguida, na etapa de pré-processamento de dados, é apresentada a normalização efetuada nos sinais de ECG. Na quarta etapa, apresenta-se as arquiteturas de CNNs propostas nesse trabalho para a detecção do complexo QRS. Na quinta etapa, realização de experimentos, descreve-se o processo de treinamento do modelo, mostrando o ajuste dos parâmetros do treinamento.

A sexta etapa é caracterizada pela avaliação do desempenho dos modelos de CNNs, com a seleção do melhor modelo. Após a escolha do melhor modelo, é

apresentada a etapa de pós-processamento, com o objetivo de remoção de ruídos indesejados presentes nos intervalos detectados pela rede CNN. Esses ruídos correspondem a falsos positivos, ou seja, outros tipos de intervalos que foram classificados como intervalos QRS. Por fim, a última etapa consiste da obtenção das métricas que avaliam o desempenho do método desenvolvido para detecção do complexo QRS, antes e depois do pós-processamento, utilizando o conjunto de teste.

Nas próximas seções desse capítulo, detalharemos cada um dos blocos anteriormente referidos.

4.1 Materiais

4.1.1 Definição do ambiente de trabalho

Para este trabalho, foi utilizado um computador com sistema operacional Windows 10, um processador Intel Core i7-8700 CPU @ 3.20GHz e 3.19GHz, 16GB de memória RAM e uma GPU NVIDIA GeForce GTX 1070 com 8GB de memória dedicada e 1920 núcleos de processamento CUDA.

Para o desenvolvimento da pesquisa, pré-processamento da base de dados e simulação das redes convolutivas, foi utilizado o software MATLAB® versão 2020b com as *toolboxes: Neural Network e Computer Vision System*.

4.1.2 Banco de Dados

O banco de dados utilizado é o *St Petersburg INCART 12-lead*, disponibilizado pelo instituto de técnicas cardiológicas de São Petersburgo na Rússia. Esse banco de dados é disponibilizado na plataforma *PhysioNet*, um dos maiores repositórios abertos de dados médicos (GOLDBERGER *et al.*, 2000)

Este banco de dados consiste em 75 gravações anotadas extraídas de 32 pacientes registrados por *Holter* (dispositivo portátil que monitora continuamente a atividade elétrica cardíaca). Cada registro tem 30 minutos de duração e contém 12 derivações padrões, cada uma amostrada a 257 Hz, com ganhos variando de 250 a 1.100 unidades do conversor analógico-digital por milivolt. Os ganhos para cada

registro são especificados em um arquivo *.hea*. Os arquivos de anotação de referência contêm ao todo mais de 175.000 anotações de batimentos cardíacos.

Os registros originais foram coletados de pacientes submetidos a exames para doença arterial coronariana (17 do sexo masculino e 15 feminino, com idades entre 18 e 80 anos; idade média: 58 anos). As informações desse banco podem ser verificadas na Tabela 1. Nenhum dos pacientes tinha marca-passo e a maioria possuía batimentos ectópicos ventriculares (ritmos cardíacos anormais intercalado aos batimentos normais). Na seleção dos registros a serem incluídos no banco de dados, foi dada preferência a indivíduos com ECGs consistentes com isquemia, doença arterial coronariana, anormalidades de condução e arritmias.

As anotações de batimentos cardíacos foram produzidas por um algoritmo automático e depois corrigidas manualmente, seguindo as definições padrão do *PhysioBank*. As derivações possuem alinhamento entre si, isto é, o início e término de uma derivação coincide com as demais.

A Figura 13 apresenta o registro I01 do paciente 1, com todas as derivações da base de dados. As anotações para cada complexo QRS são mostradas na parte superior da Figura 13. Pode-se observar que todas as derivações estão alinhadas, ou seja, fenômenos cardíacos semelhantes, em todas as derivações, ocorrem na mesma linha vertical. Para a visualização mostrada na Figura 13, foi aplicado um filtro para correção da linha de base e suavização do eletrocardiograma.

Tabela 1: Relação das informações dos pacientes presentes na base de dados INCART.

Paciente	Idade	Registro	Sexo	Diagnóstico
1	65	I01 e I02	F	Doença arterial coronariana e hipertensão arterial
2	59	I03, I04 e I05	M	Enfarte agudo do miocárdio (Ataque cardíaco)
3	80	I06 e I07	M	Ataque isquêmico transitório
4	51	I08	F	Não há
5	68	I09, I10 e I11	M	Não há
6	39	I12, I13 e I14	F	Não há
7	57	I15	M	Ataque isquêmico transitório
8	64	I16 e I17	M	Ataque isquêmico transitório
9	18	I18 e I19	F	Não há
10	59	I20, I21 e I22	F	Doença arterial coronariana, hipertensão arterial e hipertrofia ventricular esquerda.
11	52	I23 e I24	F	Infarto do miocárdio precoce
12	66	I25 e I26	F	Disfunção do nódulo sinusal
13	60	I27 e I28	M	Não há
14	41	I29, I30, I31 e I32	F	Não há
15	40	I33 e I34	M	Não há
16	38	I35, I36 e I37	M	Doença arterial coronariana, hipertensão arterial e hipertrofia ventricular esquerda.
17	60	I38 e I39	F	Doença arterial coronariana e hipertensão arterial

Paciente	Idade	Registro	Sexo	Diagnóstico
18	66	I40 e I41	M	Ataque isquêmico transitório
19	19	I42 e I43	M	Não há
20	53	I44, I45 e I46	F	Doença arterial coronariana, hipertensão arterial e hipertrofia ventricular esquerda
21	68	I47 e I48	F	Infarto do miocárdio precoce
22	70	I49 e I50	M	Não há
23	56	I51, I52 e I53	M	Não há
24	74	I54, I55 e I56	M	Infarto do miocárdio precoce
25	45	I57 e I58	F	Doença arterial coronariana e Hipertensão arterial
26	49	I59, I60 e I61	F	Não há
27	58	I62, I63 e I64	M	Enfarte agudo do miocárdio (Ataque cardíaco)
28	58	I65, I66 e I67	F	Não há
29	62	I68 e I69	M	Não há
30	64	I70 e I71	M	Não há
31	73	I72 e I73	M	Doença arterial coronariana, hipertensão arterial e hipertrofia ventricular esquerda
32	66	I74 e I75	M	Infarto do Miocárdio Precoce



Figura 13: Registro I01 do paciente 1 com as 12 derivações do ECG obtido da base de dados INCART (amplitude 5 mm/mV e velocidade de 25 mm/segundos). Na parte superior da figura observa-se as anotações feitas para complexo QRS.

4.1.3 Modelagem da base de dados

O processo de modelagem da base de dados consiste na extração das amostras e rotulagem dos intervalos do complexo QRS detectados nos ECGs da base de dados relatada no item 4.1.2. O procedimento utilizado para essa extração de amostras do registro de uma paciente consiste dos seguintes passos:

1. Para cada derivação de um paciente, 24 amostras foram selecionadas;
2. Cada amostra é constituída de 1.136 pontos amostrados do sinal original;
3. Para obter uma melhor representatividade das amostras, as 12 primeiras amostras foram selecionadas nos intervalos entre os pontos amostrados do sinal original de 1 a 13.632, enquanto as 12 amostras seguintes foram selecionadas entre os pontos amostrados do sinal original de 100.001 e 113.632.

Dessa forma, os 75 registros geraram um total de 1.800 amostras. A Figura 14 ilustra uma dessas amostras obtidas. A próxima etapa é a marcação dos pontos do complexo QRS nas 1.800 amostras na derivação II. Conforme mostrado na Figura 14, essa marcação consiste na identificação dos pontos Q, R e S, de todos os batimentos. Essa derivação foi escolhida pela facilidade na identificação dos pontos Q, R e S. Ressalta-se que, na medida em que os sinais de todas as derivações estão alinhados, a marcação do complexo QRS em uma derivação é utilizada como marcação do complexo QRS em todas as derivações.

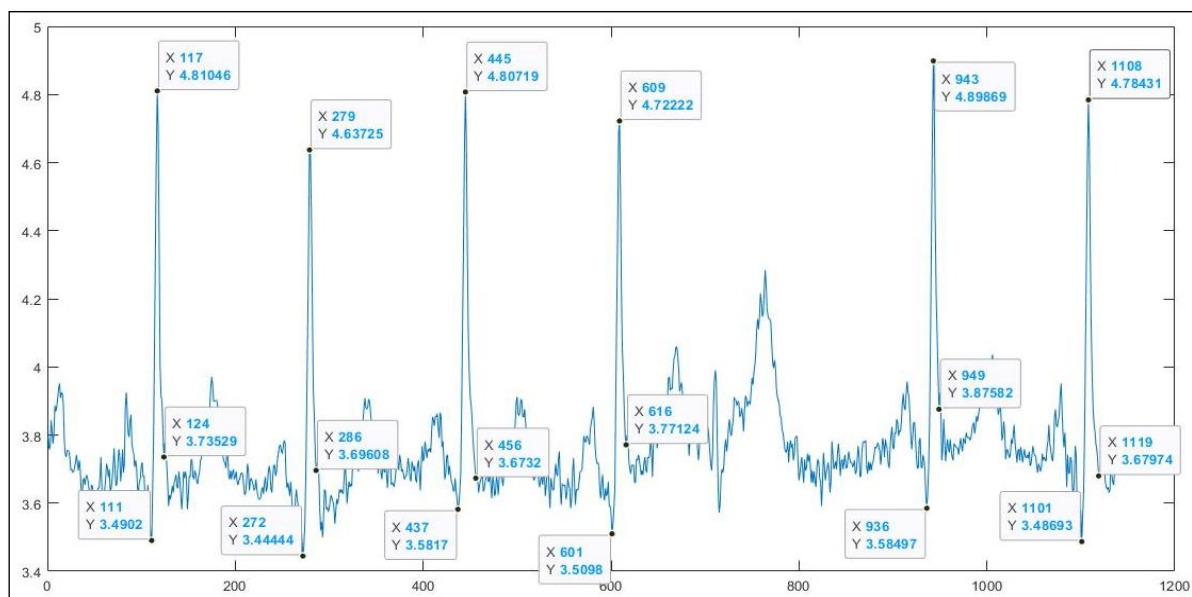


Figura 14: Ilustração da amostra 1 extraída do registro I01 do paciente 1, com identificação dos pontos Q, R e S no sinal da derivação II.

4.1.4 Geração do Sinal Padrão Ouro (*Ground Truth*)

O padrão ouro (*ground truth*) é um sinal binário, onde o nível lógico 1 (ou 255) é utilizado para demarcar os segmentos \overline{QS} , enquanto o nível lógico 0 é utilizado para marcar os outros intervalos do sinal. Um segmento \overline{QS} é identificado a partir da marcação dos pontos Q, R, S, feita na etapa anterior.

A Figura 15 mostra como é gerado o sinal utilizado como padrão ouro na saída da rede. Esse padrão ouro é utilizado tanto na etapa de treinamento como na etapa de teste das redes CNNs. A Figura 15(a) mostra a imagem de uma amostra do sinal de ECG obtido por meio do procedimento descrito no item 4.1.3, enquanto que a Figura 15(b) mostra a imagem do sinal correspondente a essa amostra. Os sinais apresentados nas Figuras 15(a) e 15(b) são exemplos de pares entrada-saída apresentados à rede na etapa de treinamento.

Ressalta-se que essa marcação é diferente da marcação feita por Koscova *et al.* (2021). Nesse trabalho, a marcação do intervalo \overline{QS} corresponde exatamente ao intervalo presente no sinal de ECG, enquanto no trabalho de Koscova *et al.* (2021), a marcação é feita do pico R, marcando-se um intervalo cujo início está 0.05ms à direita do pico R e o final está 0.05ms à direita do pico R. Esse intervalo é marcado nível lógico 1. A consequência dessa diferença de marcação, é que o treinamento realizado nesse trabalho detecta exatamente o intervalo \overline{QS} , enquanto o treinamento realizado por (KOSCOVA *et al.*, 2021) identifica apenas a localização do pico R.

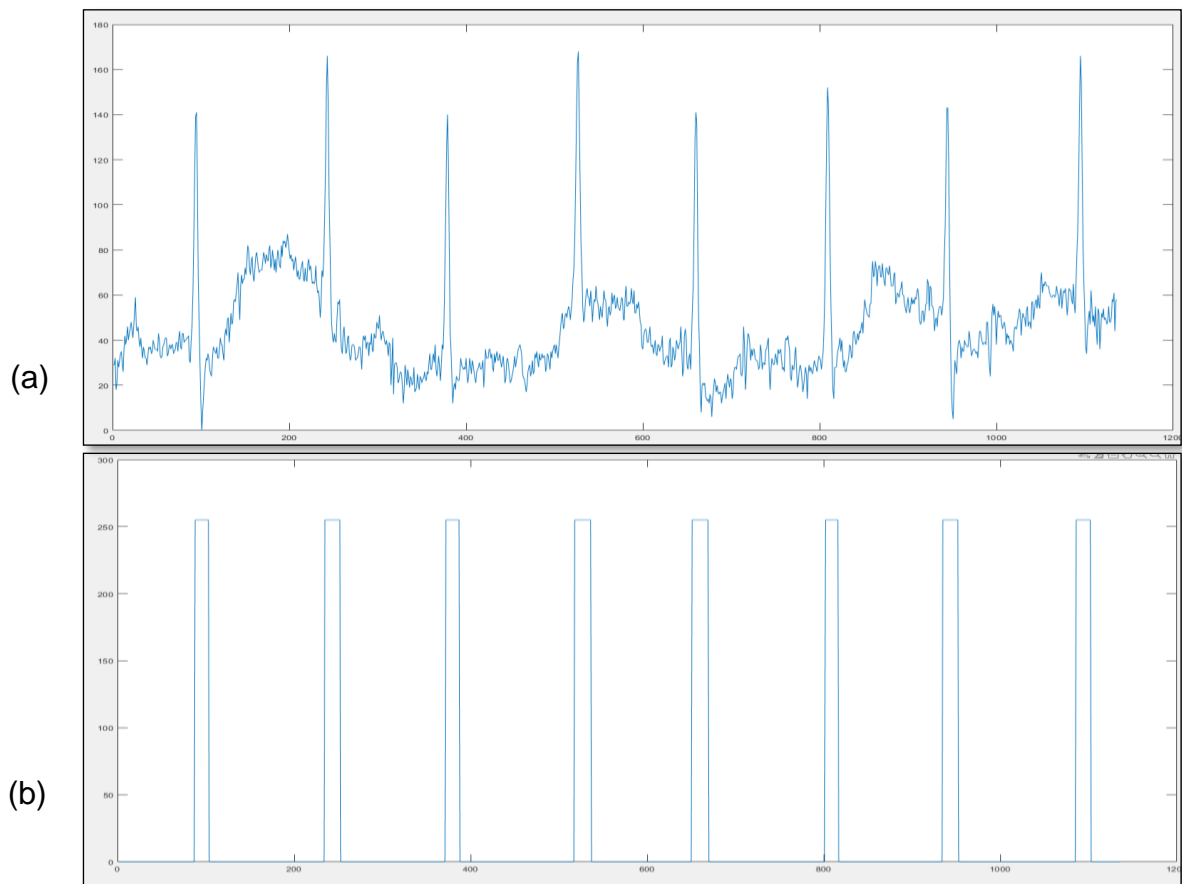


Figura 15: (a) Imagem de uma amostra de sinal extraída do sinal ECG conforme procedimento descrito no item 4.1.3; (b) padrão ouro correspondente a amostra apresentada em (a).

4.2 Métodos

4.2.1 Pré-Processamento dos dados

A normalização dos sinais de ECG é uma etapa importante para garantir a eficiência e precisão do processamento por uma rede neural convolucional (CNN).

O processo de normalização é comum em abordagens utilizando CNNs. A pesquisa de Xingjun e Qingyan (2019, p. 75), por exemplo, implementa a normalização em um sinal unidimensional de ECG com o objetivo de evitar condições anômalas na entrada da rede neural.

Neste trabalho, a amplitude do sinal ECG foi normalizada para a faixa entre 0 e 255. Este método de normalização, conhecido como normalização de escala de 8 bits, é uma abordagem comum e amplamente utilizada no processamento de imagens

digitais. Na normalização de escala de 8 bits cada valor do sinal original, I é transformado para um valor I_N , no intervalo $[0, 255]$ através da seguinte equação:

$$I_N = 255 * \left(\frac{L - 1}{I_{max} - I_{min}} \right) * (I - I_{min}) \quad (22)$$

em que:

$L = 2^8=256$, para uma imagem de 8bits;

I = Amplitude do Sinal de ECG a ser reescalonado;

I_{min} = Valor mínimo da onda Q dos complexos QRS do sinal de ECG;

I_{max} = Valor máximo da onda R dos complexos QRS do sinal de ECG;

I_N = Amplitude do sinal de ECG reescalonado;

A normalização da escala de amplitude do sinal ECG também ajuda a garantir que as camadas da CNN possam aprender as características importantes dos sinais de ECG de forma mais eficiente, sem serem influenciadas por variações na amplitude do sinal. Isso é importante, pois a amplitude dos sinais de ECG pode variar significativamente entre diferentes indivíduos ou mesmo entre diferentes registros para o mesmo indivíduo.

Para o processo de segmentação semântica, os sinais unidimensionais, constituídos por 1136 amostras, foram agrupados dois a dois, com dimensões de 2×1136 (pontos amostrados do sinal de ECG). O objetivo de agrupar os sinais unidimensionais em sinais 2D é permitir que a rede neural seja capaz de aprender padrões espaciais e realizar uma segmentação semântica, na sequência de sinais ECG. Esta abordagem também é observada no trabalho de Zhengling *et al.* (2020, p. 2). Assim como esses autores, aplicamos o agrupamento 2D para todas as derivações do sinal ECG.

4.2.2 Definição das Arquiteturas das Redes Neurais Convolucionais

Para a detecção do complexo QRS nas 12 derivações do eletrocardiograma, foram propostas três arquiteturas de redes CNNs, descritas a seguir. Estas redes são inspiradas no modelo de segmentação semântica, que classifica cada pixel em uma

imagem como pertencendo a uma dada classe. Nesse trabalho o número de classes é igual a 2. Ou seja, um pixel pertencerá a um segmento \overline{QS} , ou não.

A primeira arquitetura de rede, denotada por CNN 1, é composta por uma série de camadas, sendo que a entrada de cada camada é a saída de uma camada anterior. A primeira camada, que corresponde a camada de dados de entrada, possui dimensões de $2 \times 1136 \times 1$, ou seja, uma matriz bidimensional de comprimento 1136 que segue a abordagem citada anteriormente no item 4.2.1.

Em seguida, há três subgrupos de camadas, constituídos por: camada de convolução 2D (Conv), camada de normalização em lote (*batch normalization* - *batchnorm*), unidade retificadora linear (ReLU) e camada de amostragem (*maxpooling*).

A camada de convolução 2D é usada para realizar operações de filtragem dos sinais. Filtros unidimensionais, com formato de 1×4 (1 é a profundidade ou número de canais, e 4 é o comprimento do filtro), são aplicados para extrair recursos (características) do sinal de entrada. O segundo parâmetro da camada convolutiva é o número de filtros utilizados. No primeiro subgrupo, foram utilizados 25 filtros ou *kernels*. No segundo subgrupo foram utilizados um total de 50 filtros, enquanto que no terceiro subgrupo de camadas, foram utilizados 100 filtros. O número de filtros representa o número planos de neurônios da camada convolutiva, ou seja, sua profundidade, expressando o número de neurônios na camada convolutiva que se conectam à uma mesma região na entrada, com o mesmo tamanho do *kernel*. O parâmetro *padding* foi definido de modo a que o tamanho da imagem de entrada fosse preservado após a aplicação do filtro.

A camada de *batch normalization* normaliza a saída da camada convolutiva anterior, ajudando a acelerar o treinamento da rede neural e a evitar *overfitting*. Para fechar o bloco da unidade convolutiva temos a camada de ativação ReLU que executa uma operação de limite para cada elemento da entrada. Como resultado dessa operação, qualquer valor menor do que zero (valores negativos) assumem o valor zero, enquanto os valores positivos são reproduzidos na saída, sem alteração (função linear).

A última camada de cada subgrupo, a camada de *amostragem* 2D, reduz o tamanho espacial da representação dos dados. Os parâmetros especificados para esta camada foram o seguinte: tamanho da máscara de amostragem de 1×2 pixels

e passo (*stride*) de 1 x 2, ou seja, a cada iteração, a camada de amostragem move 1 pixel no eixo horizontal. Após amostrar os dados de uma linha da matriz, a máscara move-se 2 pixels no eixo vertical e inicia novamente a interação.

Conforme mostrado na Figura 16, a arquitetura de rede neural contém três unidades convolutivas. Em cada unidade variamos apenas o número de filtros da camada convolutiva, sendo: 25 para a primeira unidade convolutiva, 50 para a segunda e 100 para a terceira. É importante ressaltar a presença da camada de *Max pooling* após cada uma das unidades convolutivas. Após estas camadas, há uma camada de *dropout* para prevenir *overfitting* durante o treinamento.

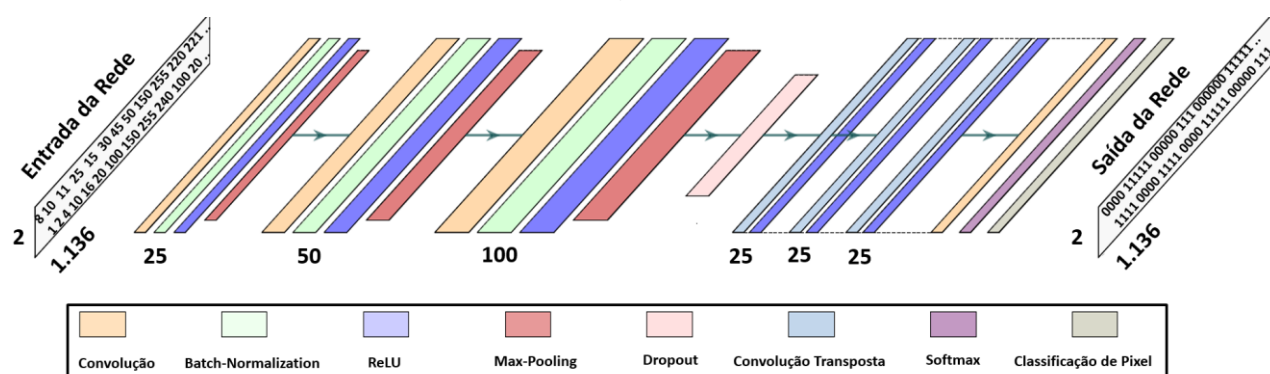


Figura 16: Representação da Arquitetura CNN1 de rede neural convolucional.

Em seguida, há três unidades transpostas sucessivas, compostas por uma camada de convolução transposta e da camada de ativação ReLU. As camadas de convolução transpostas 2D têm o objetivo de realizar a transposição da convolução, aumentando a resolução do mapa de recursos bidimensionais, ou seja, ela inverte o processo de *pooling* aplicado após as camadas de convolução. Ela retoma às estruturas iniciais características. Os parâmetros definidos para a camada de convolução transposta são: o tamanho da janela de filtro (*kernel size*) de 1x2 e o número de filtros igual a 25, sendo estes parâmetros semelhantes para todas as unidades transpostas.

Finalmente, existem mais três camadas finais da rede CNN. A primeira delas é uma camada de convolução 2D com filtro 1x2, para extração adicional de características. Em seguida temos a camada de ativação *softmax* que converte as saídas da rede CNN em probabilidades de ocorrência de cada classe. A classe mais provável é aquela que apresentar uma maior probabilidade nessa camada.

A terceira e última é a camada de classificação de pixels que produz uma classificação dos pixels para cada imagem de entrada, com base nas probabilidades fornecidas pela camada *softmax*. Essa camada usa a métrica de perda *cross-entropy* para avaliar o desempenho da rede.

A segunda, denotada por CNN2 (Figura 17-a), e a terceira, denotada por CNN3 (Figura 17-b), são arquiteturas de rede CNN que possuem as mesmas camadas da arquitetura 1, variando apenas na quantidade de filtros nas unidades convolutivas de extração de características. A arquitetura 2 possui 32 filtros na primeira unidade convolutiva, 64 na segunda e 128 na terceira, além de 32 filtros em cada uma das três unidades convolutivas transpostas. A arquitetura 3 possui 75 filtros na primeira unidade convolutiva, 150 na segunda e 300 na terceira, além de 75 filtros em cada uma das três unidades convolutivas transpostas.

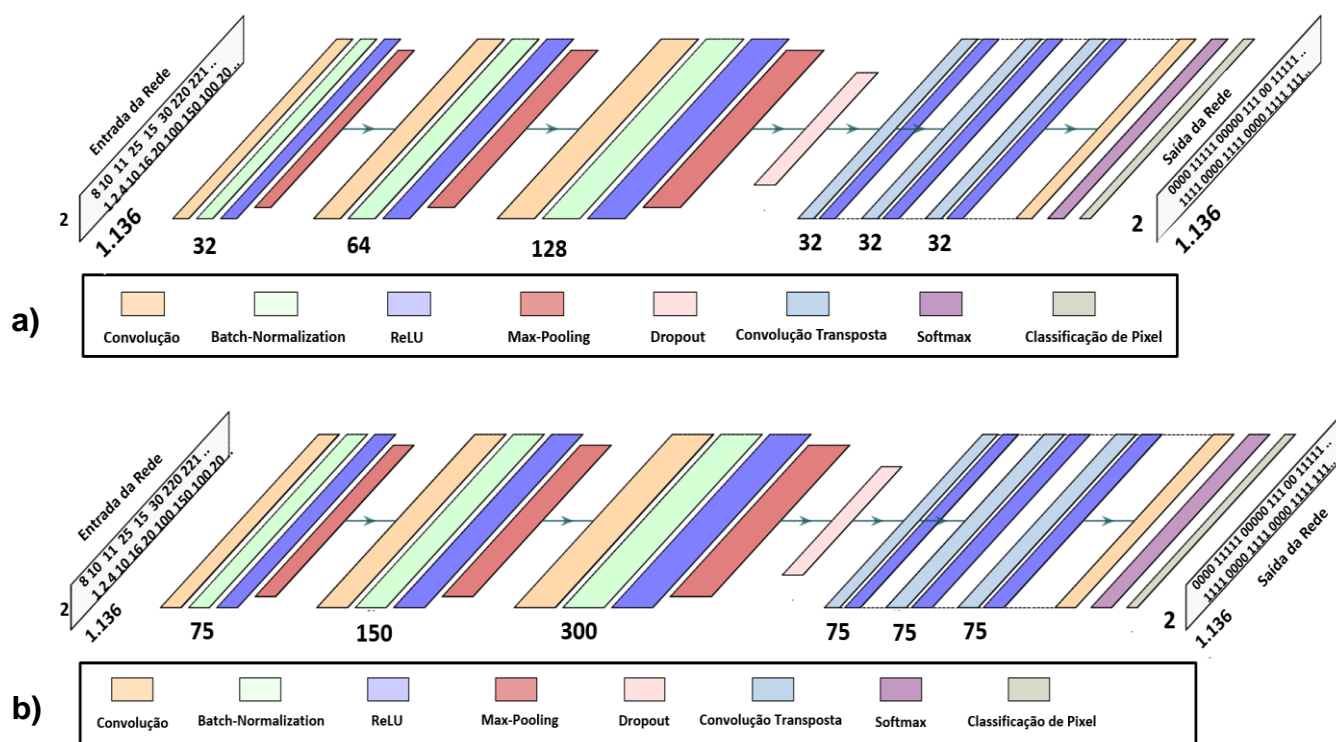


Figura 17: Representação das Arquiteturas de CNN 2 e 3 de rede neural convolucional.

A Tabela 2 resume a quantidade de filtros por unidade convolutiva para cada uma das três CNN utilizadas nesta pesquisa.

Tabela 2: Filtros por unidade convolutiva para as CNN 1, 2 e 3.

Unidade Convolutiva	Número de Filtros		
	CNN1	CNN2	CNN3
1	25	32	75
2	50	64	150
3	100	128	300

4.2.3 Experimentos

Nesta seção são apresentados os hiperparâmetros definidos para o treinamento da rede CNN, utilizada na detecção do complexo QRS nas 12 derivações do ECG. Os tópicos a serem abordados na seção são listados a seguir:

- Divisão do conjunto de dados para treinamento e teste;
- Ajuste da taxa de aprendizado;
- Ajuste dos parâmetros de treinamento;
- Procedimento de treinamento.

Após a modelagem da base de dados, descrita na seção 4.1.3, a mesma foi dividida em dois conjuntos: conjunto de treinamento, contendo 80% das amostras da base de dados, e conjunto de teste, contendo 20% das amostras da base de dados, conforme mostrado na Figura 18.

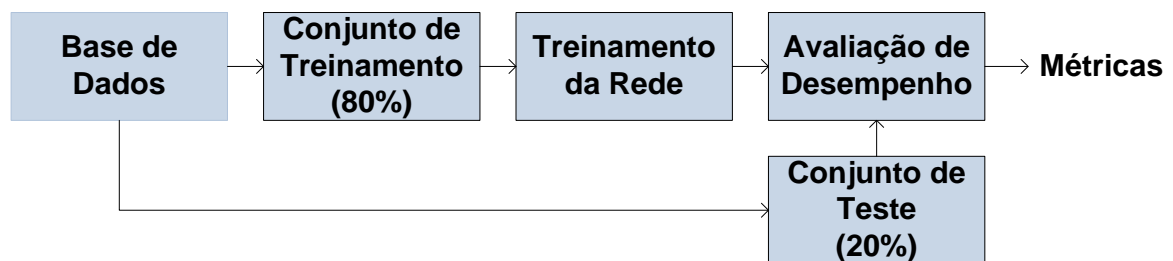


Figura 18: Divisão do conjunto de dados para ser utilizado nos experimentos com a rede neural convolucional, com 80% dos dados no conjunto de treinamento e 20% dos dados no conjunto de teste.

O método adotado para o ajuste da taxa de aprendizado consistiu em uma técnica de equilíbrio de pesos de classes. Primeiro, é utilizada uma função para contar a quantidade de pixels de cada classe presente nos dados de treinamento. Em

seguida, é calculada a frequência de cada classe como a proporção de pixels dessa classe em relação ao número total de pixels. Por fim, o peso de cada classe é calculado como o inverso da sua frequência.

Esses pesos de classe serão utilizados durante o treinamento do modelo para equilibrar a importância de cada classe e evitar problemas de desbalanceamento de classes. O objetivo é equilibrar as contribuições das diferentes classes no treinamento da rede neural, de modo que classes com menor frequência tenham pesos maiores. O cálculo do peso w para cada classe k se dá pela equação (23):

$$w_k = \frac{\text{total de pixels}}{\text{total de pixels da classe } k} \quad (23)$$

Na etapa de ajuste dos parâmetros de treinamento (Tabela 3), adotamos o uso do método de treinamento "ADAM", que é uma das abordagens mais populares e eficientes para otimização de redes neurais. Além disso, utilizamos algumas opções específicas para o treinamento, como a taxa inicial de aprendizado de 0,0003, fator de decaimento do gradiente ao quadrado de 0,99, número máximo de épocas de 300, *mini-batch* de 64 e embaralhamento dos dados a cada época do treinamento

Tabela 3: Representação dos parâmetros para o treinamento do modelo.

Parâmetros	Ajustes
Algoritmo de treinamento	ADAM
Taxa de Aprendizado Inicial	0,0003
Fator de queda da taxa de aprendizagem	0,99
Período de queda da taxa de aprendizagem	300 épocas
Tamanho do lote	64

Essas opções de hiperparâmetros foram definidas com o objetivo de garantir uma convergência mais rápida e precisa do modelo. No capítulo de resultados, ao mostramos as curvas de acurácia e perda, versus o número de épocas, veremos que as escolhas anteriormente feitas foram acertadas, pois veremos uma convergência muito bem comportada no treinamento do modelo.

No procedimento de treinamento, as 12 derivações foram treinadas individualmente utilizando as arquiteturas propostas no item 4.2.2, resultando em 36

experimentos (12 derivações x 3 arquiteturas de rede). A entrada da rede foi adaptada para que em cada experimento a derivação a ser utilizada estivesse de acordo com o modelo treinado. A Figura 19 exemplifica o processo de treinamento com as 12 derivações utilizando a arquitetura 1. O mesmo procedimento é aplicado no treinamento das arquiteturas 2 e 3.

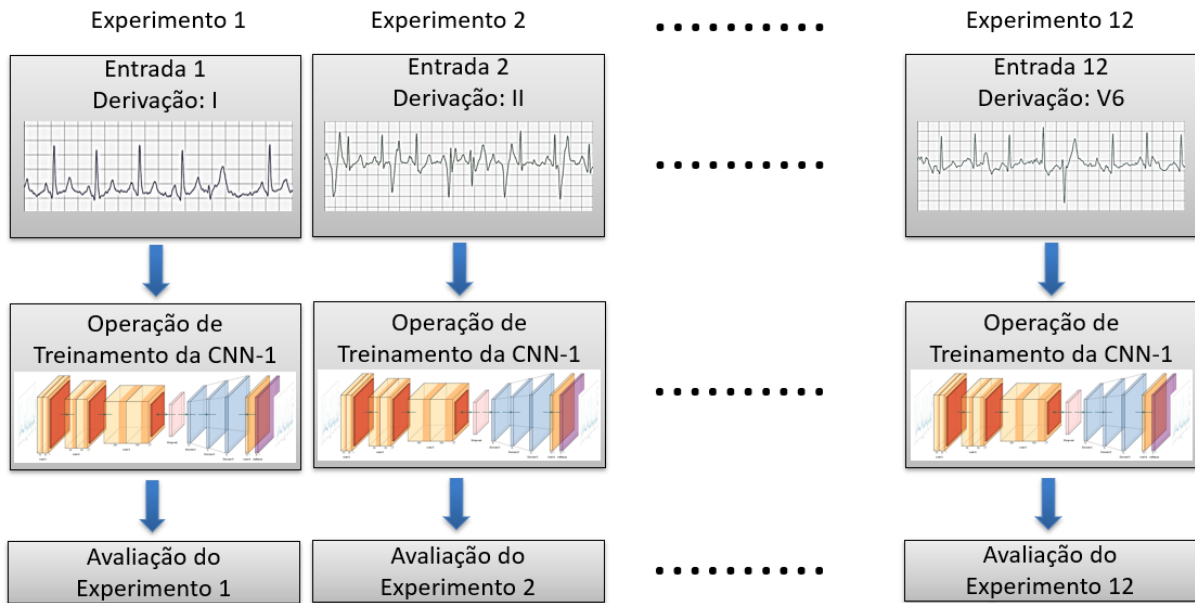


Figura 19: Procedimento de treinamento para as 12 derivações do ECG.

4.2.4 Pós-Processamento

O procedimento de pós-processamento garante a retirada de falsos positivos (intervalos não QRS) do segmento ECG analisado. Esses falsos positivos são gerados, às vezes, devido a ondas P ou artefatos no sinal. Os passos a seguir demonstram o procedimento adotado para remoção de falso positivo gerado por uma onda P:

Passo-1: Utilizar o resultado da rede neural como uma máscara para filtrar na imagem de entrada os intervalos identificados, conforme a Figura 20:

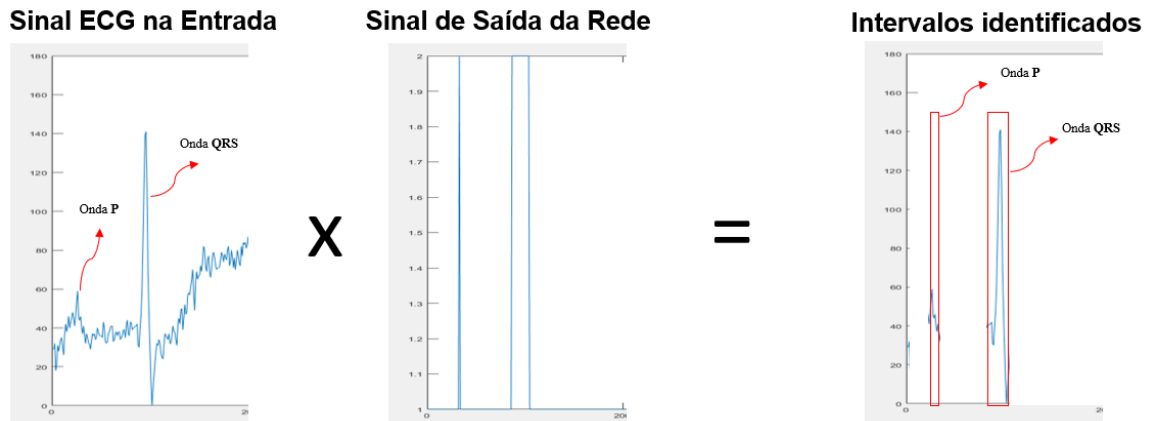


Figura 20: Passo 1 da etapa de pós-processamento: delimitação no sinal original dos intervalos \overline{QS} identificados pela rede neural.

- **Passo-2:** Nos intervalos identificados são obtidos os pontos de máximo e mínimo, possibilitando calcular a altura do intervalo fazendo a diferença entre os extremos. A Figura 21 apresenta o resultado do passo 2.

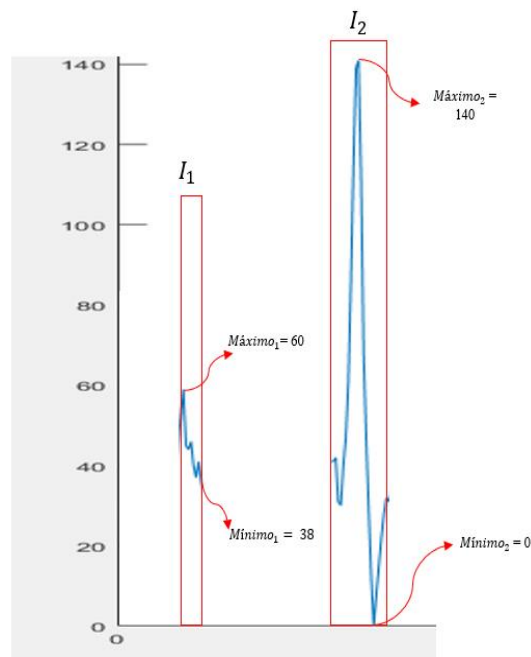


Figura 21: Passo 2 da etapa de pós-processamento: identificação dos pontos de máximos e mínimos para o cálculo da amplitude do sinal em cada intervalo \overline{QS} detectado.

$I = \{I_1, I_2, \dots, I_n\}$ representa os intervalos identificados em uma amostra j de ECG e $H = \{H_1, H_2, \dots, H_n\}$ refere-se a altura para cada intervalo, sendo calculada pela relação:

$$H_i = |\text{Máximo}_i - \text{Mínimo}_i| \quad (24)$$

Para o exemplo, temos:

$$\begin{aligned} H_1 &= |60 - 38| = 22 \\ H_2 &= |140 - 0| = 140 \end{aligned} \tag{25}$$

- **Passo-3:** Para cada amostra j , obter \bar{A}_j , que representa a média das amostras no intervalo i .

$$\bar{A}_j = \frac{\sum_{i=1}^n H_i}{n} \tag{26}$$

Em que n é o número de intervalos \overline{QS} detectados na amostra j .

- **Passo-4:** Aplicar a seguinte regra para filtrar os ruídos

Se intervalo $H_i > \frac{\bar{A}_j}{2}$

Então I_i corresponde a um intervalo \overline{QS} com uma onda QRS

Senão não corresponde a um intervalo \overline{QS} (falso positivo detectado pela CNN)

4.2.5 Métricas de Desempenho

A seguir, definiremos as métricas utilizadas para avaliar o desempenho das Redes Neurais Convolucionais (CNN). Essas métricas não apenas comparam o desempenho das arquiteturas propostas neste trabalho, mas também o desempenho dessas arquiteturas com outros resultados obtidos na literatura. A seguinte terminologia é usada para definir as métricas usadas abaixo:

- **TP:** Verdadeiros Positivos (caso positivo previsto corretamente como positivo);
- **TN:** Verdadeiros Negativos (caso negativo previsto corretamente como negativo);
- **FP:** Falso Positivo (caso negativo predito incorretamente como positivo);
- **FN:** Falso Negativo (caso positivo previsto incorretamente como negativo).

Temos um verdadeiro positivo (TP) quando um *pixel* da classe QRS é corretamente classificado pelo preditor. Ocorre um falso positivo (FP) se um *pixel* da classe não QRS é classificado de forma incorreta como QRS. Ocorre um verdadeiro negativo (TN) caso um *pixel* da classe não QRS seja corretamente classificado como tal e, por último, ocorre um falso negativo (FN) se um *pixel* da classe QRS é classificado incorretamente como não QRS.

Assim como Liu *et al.* (2020, p. 416) e outras pesquisas da literatura, como a de Zhong, Liao e Wang (2018), adotamos a precisão, o *recall* e o f1-score como métricas

para avaliar o desempenho da detecção do complexo QRS em eletrocardiografias. A seguir definimos essas métricas:

1) Precisão

A métrica precisão avalia a capacidade da rede neural de identificar corretamente quais pixels são pertencentes ou não ao complexo QRS. A precisão avalia a quantidade de segmentos que, dentre todos os segmentos classificados como positivos, são da classe positiva, conforme a equação (27).

$$Precisão = \frac{TP}{TP + FP} \quad (27)$$

Uma precisão alta indica que a maioria dos pixels classificados como positivos são realmente positivos, enquanto uma precisão baixa indica que existe uma grande quantidade de pixels negativos classificados como positivos.

2) Recall

A métrica *recall* (também conhecida como taxa de verdadeiros positivos ou sensibilidade) é uma medida de desempenho de um classificador que avalia a proporção de exemplos positivos corretamente identificados em relação ao total de exemplos positivos. O *recall* mede a capacidade do classificador em identificar corretamente todas as ocorrências de uma classe específica.

$$Recall = \frac{TP}{TP + FN} \quad (28)$$

De acordo com a equação (28) o *recall* é calculado como a razão entre o número de verdadeiros positivos e a soma do número de verdadeiros positivos e falsos negativos. Quanto maior o *recall*, melhor é a capacidade do classificador em identificar todas as ocorrências da classe positiva.

3) F1-score

A *F1-score* é uma medida harmônica que equilibra a precisão e o *recall* e, portanto, pode ser uma medida mais representativa do desempenho geral da

detecção do complexo QRS. Além disso, a *F1-score* é especialmente útil em casos em que o número de verdadeiros positivos é pequeno em relação ao número total de instâncias, como é comumente o caso na detecção do complexo QRS. A equação (29) foi adotada para o cálculo da métrica de desempenho *F1-score*.

$$F1\ score = \frac{2 \times Precisão \times Recall}{Precisão + Recall} \quad (29)$$

5 RESULTADOS E DISCUSSÕES

Nesta seção, são apresentados os resultados dos experimentos da etapa de treinamento e teste da arquitetura CNN1, 2 e 3 para as 12 derivações do sinal de ECG. Ressaltamos que, diferentemente de outros métodos publicados na literatura, o método ora proposto identifica não apenas a presença de um complexo QRS, mas também delimita o início e o fim do intervalo \overline{QS} .

5.1 Convergência da Rede Neural Convolutacional

A convergência do algoritmo de treinamento das arquiteturas CNN1, 2 e 3, durante a etapa de treinamento, é apresentada através das curvas da acurácia e da perda em função do número de iterações. A Figura 22 mostra essas curvas com respeito aos dados de treinamento da arquitetura CNN1, com os dados da derivação I.

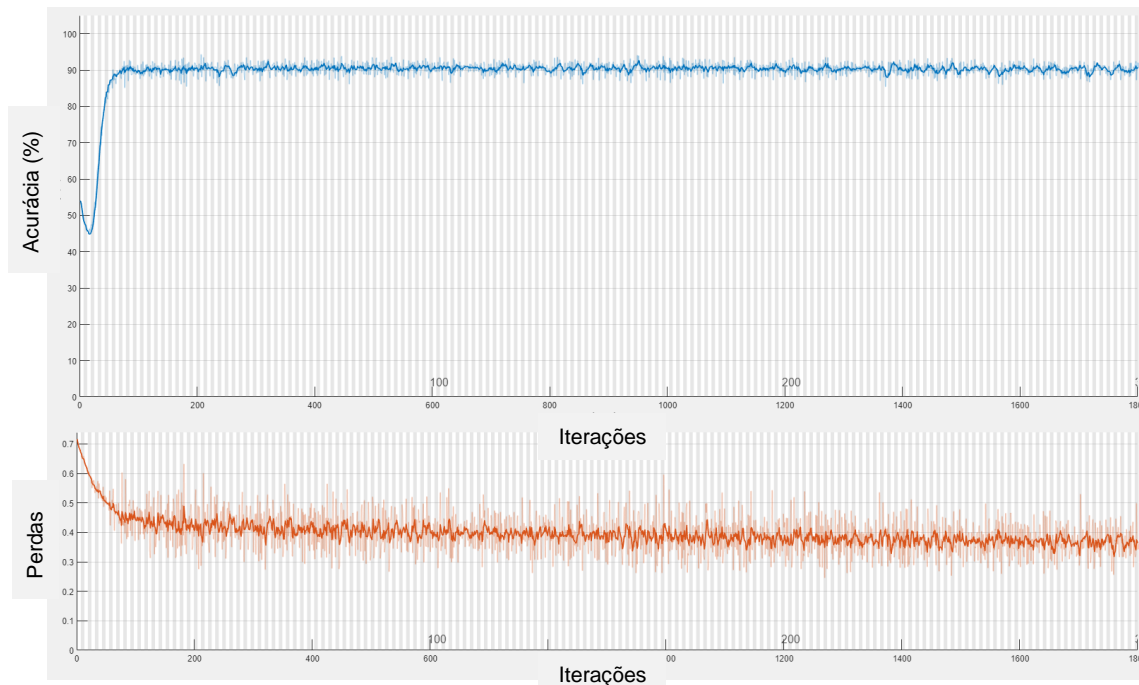


Figura 22: Gráfico de convergência das curvas de acurácia e perda em função do número de iterações do treinamento, para a arquitetura CNN1. Foram utilizados dados da derivação I.

O treinamento das redes CNN1, 2 e 3, nas 12 derivações, ocorreu em 1800 interações e 300 épocas, com taxas de acurácia e perda constantes ao longo das

últimas épocas do treinamento. Um fator de diferenciação observado no treinamento das diversas derivações foi o tempo de treinamento. Um tempo maior, cerca de 12 minutos, foi observado no treinamento das seguintes derivações: V2, V3, V4 e V5. Este fato está relacionado com as características do sinal nestas derivações, que apresentam trechos ruidosos. No entanto, a presença desses ruídos não implicou numa diminuição da acurácia, nem no aumento da perda.

5.2 Resultados de Testes dos Modelos

Os resultados utilizando as arquiteturas de CNN1, 2 e 3 para o conjunto de teste, são mostrados a seguir. A Tabela 4 apresenta os resultados das métricas de Precisão, *Recall* e F1-score usados para avaliar o desempenho dos modelos antes e depois da aplicação do pós-processamento. As linhas da Tabela 4 representam cada uma das 12 derivações do ECG testadas separadamente e nas colunas há as métricas de desempenho para cada um dos modelos.

Analisando os resultados médios das 12 derivações mostrados na Tabela 4, observa-se que, antes do pós-processamento, é possível verificar que a menor precisão, 67,55%, é obtida com a CNN3 e a maior, com a CNN2, alcançando o valor de 76,44%. Já a taxa de *recall*, para todos os modelos, está acima de 99%. Isso significa que o modelo tem uma boa capacidade de identificar corretamente as amostras com segmento \overline{QS} (*recall* alto).

Interpretando de outra forma, poderíamos afirmar que as CNNs estão classificando muitos casos como positivos (segmento \overline{QS}), mesmo que eles não o sejam, resultando em uma grande quantidade de falsos positivos.

A aplicação do pós-processamento proporcionou uma diminuição dos falsos positivos em todas as derivações das três CNNs. Com o pós-processamento, o menor valor médio observado para a precisão, nas 12 derivações, foi de 99,96%, com a CNN3. A CNN2 apresentou o maior valor médio para a precisão, após o pós-processamento, chegando a 99,98%, ou seja, o método proposto foi capaz de eliminar os falsos positivos em praticamente todos os modelos.

Tabela 4: Os resultados das métricas de precisão, *recall* e *F1-score* para a arquitetura CNN1, 2 e 3, no conjunto de teste, para as 12 derivações, antes e depois do pós-processamento.

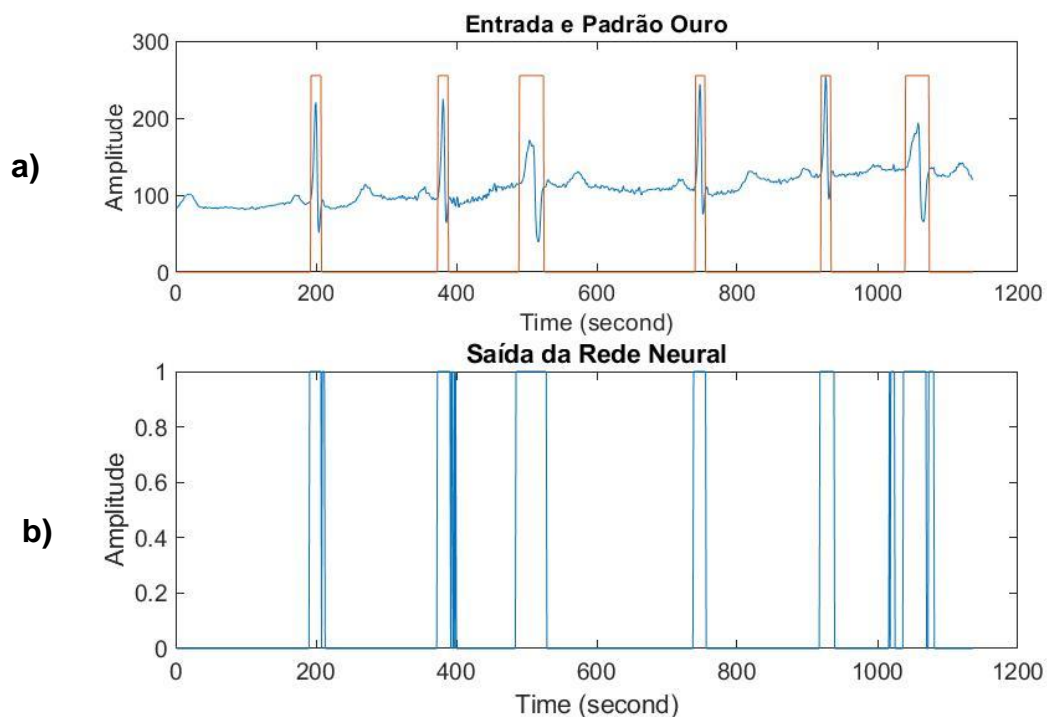
Derivação	CNN1						CNN2						CNN3					
	Antes do Pós-Processamento			Depois do Pós-Processamento			Antes do Pós-Processamento			Depois do Pós-Processamento			Antes do Pós-Processamento			Depois do Pós-Processamento		
	Prec. (%)	Recall (%)	F1-score (%)	Prec. (%)	Recall (%)	F1-score (%)	Prec. (%)	Recall (%)	F1-score (%)	Prec. (%)	Recall (%)	F1-score (%)	Prec. (%)	Recall (%)	F1-score (%)	Prec. (%)	Recall (%)	F1-score (%)
I	66,22	99,73	79,59	99,96	99,73	99,85	75,03	99,79	85,66	99,95	99,79	99,87	66,12	99,77	79,53	99,92	99,77	99,85
II	64,39	100	78,34	100	100	100	71,30	100	83,25	100	100	100	64,66	100	78,54	99,96	100	99,98
III	69,24	100	81,82	100	100	100	71,38	100	83,3	100	100	100	69,05	100	81,69	100	100	100
aVR	79,00	99,94	88,25	99,83	99,94	99,89	83,13	100	90,79	99,94	100	99,97	79,7	99,94	88,68	99,77	99,94	99,85
aVL	61,34	99,56	75,91	99,94	99,56	99,75	72,24	99,22	83,61	99,95	99,22	99,59	61,34	99,53	75,9	99,97	99,53	99,75
aVF	71,59	100	83,45	100	100	100	79,52	100	88,59	100	100	100	71,18	100	83,16	100	100	100
V1	69,85	100	82,25	100	100	100	76,50	100	86,69	100	100	100	68,92	100	81,6	100	100	100
V2	75,29	99,68	85,79	100	99,68	99,84	78,77	99,83	88,06	100	99,83	99,91	74,95	99,48	85,49	100	99,48	99,74
V3	63,94	100	78,01	99,97	100	99,98	72,67	100	84,17	99,95	100	99,98	63,24	100	77,48	100	100	100
V4	66,38	99,96	79,78	99,92	99,96	99,94	69,96	99,96	82,31	99,96	99,96	99,96	66,75	99,96	80,05	99,96	99,96	99,96
V5	60,88	100	75,68	100	100	100	75,90	100	86,3	100	100	100	60,99	100	75,77	100	100	100
V6	63,62	100	77,77	100	100	100	90,90	100	95,23	100	100	100	63,73	100	77,85	99,97	100	99,98
Média	67,65	99,91	80,55	99,97	99,91	99,94	76,44	99,90	86,50	99,98	99,90	99,94	67,55	99,89	80,48	99,96	99,89	99,93

Para demonstração visual dos experimentos, selecionamos amostras de dois pacientes. O primeiro paciente foi o de número 20. Para o registro I44 desse paciente, selecionou-se a primeira amostra para análise. Esse paciente possui histórico de doença arterial coronariana, hipertensão arterial e hipertrofia ventricular esquerda.

Na hipertrofia ventricular esquerda, o coração torna-se espesso e mais largo do que o normal. Isso pode afetar a forma como o sinal elétrico se move através do coração, resultando em alterações no sinal de ECG. Dentre estas alterações pode haver um intervalo \overline{QS} ampliado, a onda R pode ser mais larga e mais alta, o complexo QRS pode estar deslocado para a direita e o intervalo \overline{QT} pode ser prolongado devido a alterações nas propriedades de condução do coração.

Na Figura 23(a), apresentamos o sinal da amostra selecionada em azul e o *ground truth* em vermelho, sobreposto ao sinal de entrada. Na Figura 23(b) mostramos o resultado da saída da rede neural com alguns falsos positivos. Para a remoção dos falsos positivos detectados, aplicou-se o pós-processamento. Na Figura 23(c), os intervalos \overline{QS} detectados, após o pós-processamento, são apresentados.

Paciente: 20, Registro: I44, Vetor: 74, Derivação: 1



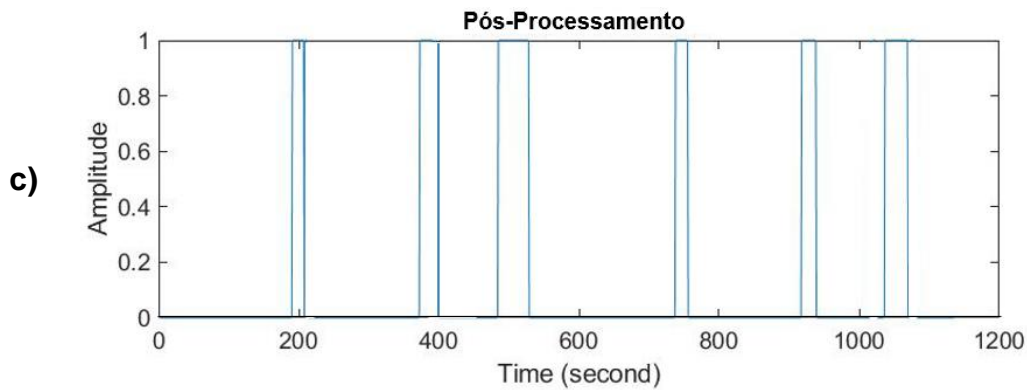


Figura 23: Resultado antes e depois do pós-processamento na amostra do ECG do paciente 20: (a) sinal original de ECG com o padrão ouro sobreposto; (b) intervalos \overline{QS} detectados pela rede convolucional; (c) intervalos \overline{QS} detectados após a aplicação do pós-processamento.

O segundo paciente selecionado foi o de número 21. Para o registro I47 desse paciente, selecionou-se a primeira amostra para análise).

Este paciente possui histórico de infarto do miocárdio precoce. O infarto do miocárdio precoce é uma condição em que ocorre o bloqueio súbito de uma ou mais artérias coronárias, interrompendo o suprimento de sangue e oxigênio para uma parte do músculo cardíaco. Isso pode resultar em lesões permanentes no miocárdio e, se não for tratado rapidamente, pode levar à insuficiência cardíaca ou à morte súbita. No ECG, é possível observar alterações nas ondas e intervalos, como a elevação do intervalo \overline{ST} e a inversão da onda T, que são indicativas de infarto do miocárdio.

Paciente: 21, Registro: I47, Vetor: 1106

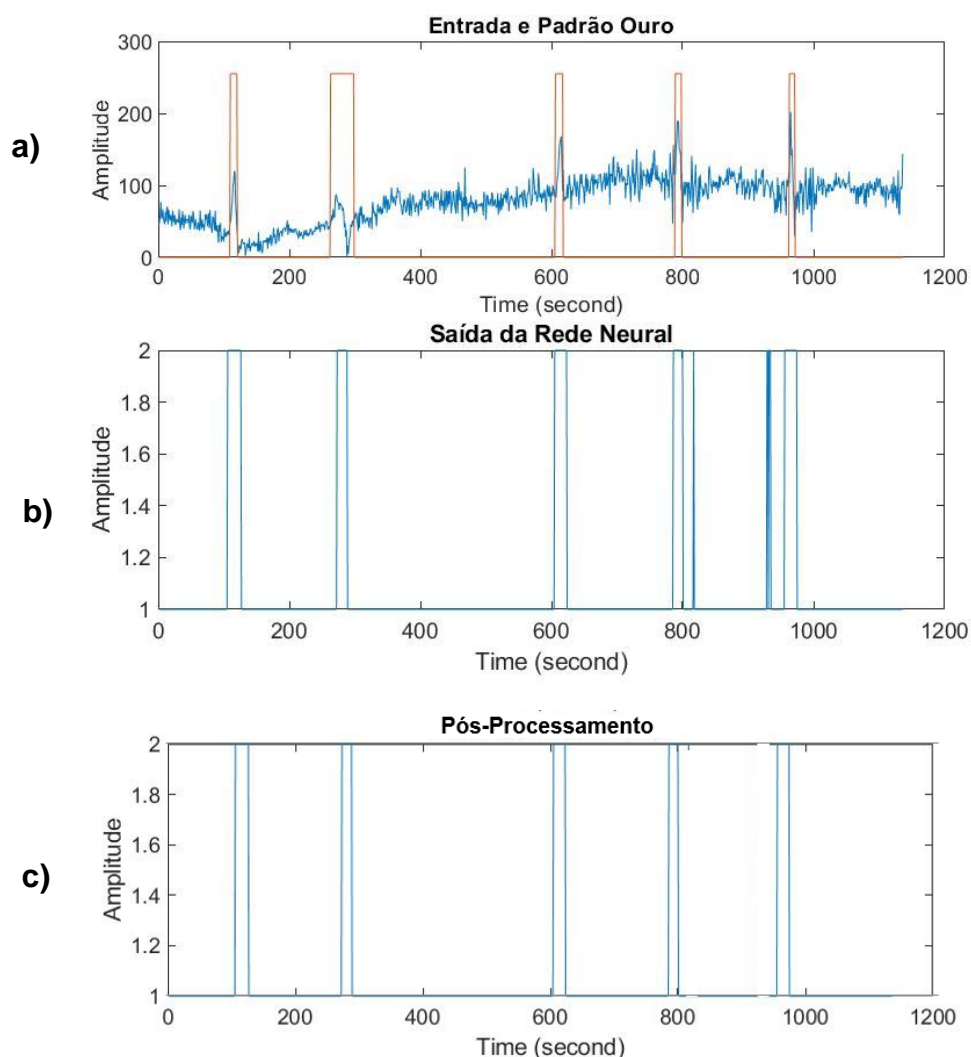


Figura 24: Resultado antes e depois do pós-processamento nas amostras do paciente 21. (a) Sinal original de ECG com o padrão ouro sobreposto; (b) intervalos \overline{QS} detectados pela rede convolucional; (c) intervalos \overline{QS} detectados após a aplicação do pós-processamento.

Na Figura 24(a), apresentamos o sinal da amostra selecionada em azul e, sobreposto, o padrão ouro em vermelho. Na Figura 24(b) mostramos o resultado da saída da rede neural, com alguns falsos positivos detectados. Para a remoção dos falsos positivos aplicou-se o pós-processamento. Na Figura 24(c) mostramos os intervalos \overline{QS} detectados, após o pós-processamento. Nela pode-se notar a retirada dos intervalos de falsos positivos detectados.

No sinal mostrado na Figura 24(a) observamos ruído e o não alinhamento da linha de base, que apresenta um comportamento oscilatório. Porém, mesmo nessas ocorrências a rede neural consegue convergência nos intervalos \overline{QS} , em consonância com o padrão ouro. Quanto à largura dos intervalos \overline{QS} percebe-se que alguns

intervalos houve uma diminuição, mas a posição das ondas *Q*, *R* e *S* se mantém detectado.

5.3 Discussão dos resultados

Os dados apresentados na Tabela 4 buscam contrastar as três arquiteturas de redes convolucionais CNN1, CNN2 e CNN3, com o objetivo de identificar qual seria a mais recomendada para a detecção do complexo QRS em 12 derivações do eletrocardiograma.

A métrica de *F1-score* demonstra uma média harmônica entre a precisão e o *recall*. Antes do pós-processamento, o maior valor médio para essa métrica, 86,50%, foi obtido com a CNN2. Com a CNN1, o valor médio obtido foi de 80,55%. Com a CNN3, o valor médio obtido foi de 80,48%. A explicação para esses valores pode ser buscada no tamanho das 3 arquiteturas utilizadas. Para extração de características, a rede CNN1 utiliza apenas 25 filtros nas camadas convolutivas. Já a CNN2 utiliza 32 filtros e a CNN3, 75 filtros. Supomos que na rede CNN1 esteja ocorrendo *underfitting*, ou seja, o número de campos receptivos utilizados para extração das características não foram necessários para identificação do complexo QRS. Supomos que também que na rede CNN3 esteja ocorrendo *overfitting*, ou seja, um número muito grande de filtros nas camadas convolutivas está fazendo com que a rede “decore” os padrões do conjunto de treinamento, não aprendendo as características mais gerais dos dados. O melhor número de filtros foi, então, aquele utilizado na CNN2, que permitiu uma boa generalização da rede no conjunto de teste.

A Tabela 5 apresenta uma análise comparativa entre os principais trabalhos que utilizam de redes neurais convolutivas para a tarefa de detecção do complexo QRS nas 12 derivações do ECG. A base de dados comum entre os autores é a INCART, a mesma utilizada nesta pesquisa, mas apenas os autores de Koscova, Z. et. al. (2021) optaram por utilizar uma base de dados própria.

Os métodos investigados utilizaram para o treinamento de seus modelos os dados das derivações juntas, logo o resultado apresentado não demonstra o desempenho do modelo em cada derivação específica. É mister frisar, que no trabalho ora apresentado, a detecção do complexo QRS é feita em cada uma das 12 derivações do ECG, o que permitiu obter métricas individualizadas para cada uma

delas. Para haver uma comparação do desempenho desta abordagem com os demais autores foi obtido a média das 12 derivações para as métricas de avaliação.

Observando as métricas de precisão, *recall* e *F1-score*, é possível perceber que o modelo proposto nesta pesquisa e o trabalho de Xiang et. al. (2018), nas primeira e segunda linhas da Tabela 5, apresentaram resultados muito próximos, com valores acima de 99% em todas as métricas. Ambos os modelos utilizaram a arquitetura CNN, mas o modelo proposto nesta pesquisa utilizou uma arquitetura de convolução transposta dedicada a segmentação semântica do segmento \overline{QS} , enquanto o modelo de Xiang et. al. (2018) propõe um sistema de extração de característica baseado em atenção composta por CNN1-2D, que extrai características morfológicas em diferentes granularidades do sinal ECG e em seguida são concatenados e enviados para um MLP, que detecta o complexo QRS.

O modelo *Faster R-CNN*, utilizado por Yu et al. (2018), apresentou um desempenho ligeiramente inferior aos modelos da primeira e segunda linha, mas ainda assim apresentou resultados muito bons, com um *F1-score* de 98,66%. Nesta abordagem o autor utiliza como entrada da rede neural o sinal na forma de imagens, logo o intervalo do complexo QRS é detectado utilizando regiões de interesses na imagem.

O modelo proposto por Liu et al. (2020) apresentou um desempenho inferior em relação às métricas de precisão, *recall* e *F1-score*, mas ainda assim apresentou resultados razoáveis, com um *F1-score* de 95,68%. Esse modelo utilizou apenas CNN para a detecção do complexo QRS. Outro aspecto importante a ser notado nesta pesquisa é o uso da etapa de pós-processamento, onde aplica-se um limiar de 0,5 para classificação binária do complexo QRS no intervalo de tempo. Assim como proposto na pesquisa de Liu et al. (2020), a etapa de pós-processamento utilizado nesta pesquisa aplica um limiar que considera a média das alturas do pico R.

O modelo proposto por Habib et al. (2021) apresentou apenas a métrica de *F1-score*, que foi de 91,00%, sem informar as métricas de precisão e *recall*. Nesta pesquisa, é formulado duas tarefas a primeira consiste no uso de uma rede neural convolutiva que irá fazer a predição do complexo QRS, a segunda utiliza um modelo de aprendizado de máquina chamado Unidades Recorrentes Fechadas do inglês *Gated Recurrent Unit* (GRU, é um tipo de rede neural recorrente como uma memória

de longo prazo com um registro de esquecimento) para aprender os limiares dos picos R no pós-processamento automaticamente, sem a necessidade de ajuste manual.

Já o modelo proposto por Koscova *et al.* (2021) apresentou resultados um pouco inferiores em todas as métricas, com um *F1-score* de 94,01%. Esta proposta utiliza apenas 3 derivações de uma base de dados com 12 derivações. Estes dados foram utilizados para o treinamento de uma rede neural convolutiva com arquitetura do tipo U-Net. A etapa de pós-processamento adotada por Koscova *et al.* (2021) refere-se a uma classificação binária das probabilidades de classes em cada amostra. O autor somente considera amostras com probabilidades acima do limiar de 0,955 e que sejam mais longas que 50 amostras ou 0,01 segundos.

Tabela 5: Análise comparativa dos trabalhos com propostas semelhantes utilizando base de dados com 12 derivações.

Trabalhos	Modelos	Base de Dados	Prec. (%)	Recall (%)	F1-score (%)
Método Proposto	CNN	INCART	99,98	99,90	99,94
Xiang <i>et. al.</i> (2018)	CNN + MLP	INCART	99,89	99,86	99,87
Yu, R. <i>et. al.</i> (2018)	<i>Faster</i> R-CNN	INCART	99,01	98,32	98,66
Liu, W. <i>et. al.</i> (2020)	CNN	INCART	96,37	94,99	95,68
Habib <i>et. al.</i> (2021)	CNN + RNN	INCART	-	-	91,00
Koscova, Z. <i>et. al.</i> (2021)	CNN: U-Net	Própria	95,00	94,89	94,01

O método proposto demonstra a utilização de uma abordagem inovadora para a detecção do complexo QRS em uma base de dados com 12 derivações. Além de desenvolver um método próprio de rotulagem da base de dados, o que contribuiu para a validação dos resultados obtidos. A utilização da base de dados INCART, comumente utilizada por outros autores, também é um ponto positivo, pois permite a comparação dos resultados obtidos com outros estudos.

A utilização de redes neurais convolutivas com convolução transposta para a segmentação semântica é uma proposta diferenciada que permitiu a obtenção do intervalo exato de início e término do segmento \overline{QS} em todas as derivações. Observe, porém, que, em intervalos onde há a inversão da onda T, como nos casos de infarto do miocárdio, o que torna o segmento \overline{QS} mais largo, o modelo sempre estimou um intervalo \overline{QS} menor. Este fato deve-se a baixa quantidade de dados com esta anomalia apresentado ao modelo.

Por fim, destaca-se o uso de uma abordagem de pós-processamento robusta que elimina os falsos positivos detectados pelo modelo. Isso contribui para a confiabilidade dos resultados e aumenta a utilidade prática da pesquisa, já que a detecção incorreta de QRS pode levar a diagnósticos equivocados e incorretos.

A associação de redes convolutivas, realizando a segmentação semântica do segmento \overline{QS} , com um estágio de pós-processamento, baseado na avaliação da altura da amplitude das ondas presentes no sinal de ECG, permitiu que se atingisse o estado da arte na detecção do complexo QRS.

A Tabela 6 apresenta um levantamento das principais características e abordagens encontradas no estado da arte em comparação com o método proposto. Ressalta-se que todas as abordagens fazem uso de redes neurais convolutivas no processo de detecção do complexo QRS. O método proposto destaca-se em comparação com os demais métodos por fazer a detecção do complexo QRS nas 12 derivações individualmente, enquanto os demais autores (PARALIC, 2019; YUEN *et al*, 2019; HE *et al*, 2020; YUEN *et al*, 2020; LIU *et al*, 2020 e HABIB *et al*, 2020) utilizam uma ou algumas poucas derivações. Todos os autores, com exceção de Koscova *et al* (2021), utilizaram bases de dados de repositórios públicos e que continham a rotulagem de marcação do pico R. Ressaltamos que no trabalho ora proposto a rotulagem foi efetuada de forma diferente. Em todas as derivações, as amostras que estão contidas em um intervalo \overline{QS} de um complexo QRS são rotuladas com 1, enquanto as outras amostras são rotuladas com zero. Assim, o que se detecta nesse trabalho é a ocorrência de um complexo QRS e a extensão do intervalo \overline{QS} , sem a necessidade de pré-processamentos especiais baseados em filtros ou operações de diferença. Quanto a arquitetura do modelo, a sequência de camadas convolutivas permitiu extrair características espaciais das diferentes ondas presentes no sinal de ECG, permitindo a diferenciação entre um complexo QRS de outras ondas. .

Tabela 6: Análise comparativa entre o estado da arte e o método proposto, comparação entre abordagens que utilizam modelos baseados em redes neurais convolutivas.

Características	Abordagens	XIANG, Yande et al. (2018)	YU, Runze et al. (2018)	PARALIC, Martin. (2019)	YUEN, Brosnan et al. (2019)	HE, Runnan et al. (2020)	YUEN, Brosnan et al. (2020)	LIU, Wei et al. (2020)	HABIB, Ahsan et al. (2021)	KOSCOVA, Zuzana et al. (2021)	DURAJ, Konrad et al. (2022)	Método Proposto
		Objetivo	Detectar o complexo QRS nas 12 derivações individualmente									
	Detectar o complexo QRS com todas as 12 derivações juntas		x							x	x	
	Detectar o complexo QRS em menos de 12 derivações	x		x	x	x	x	x	x			
	Segmentação semântica do ECG em 12 derivações (QRS, onda T, P e fundo)										x	
Base de dados	INCART com 12 derivações	x	x					x				x
	LUDB ou Próprio conjunto de dados contendo 12 derivações.									x	x	
	MIT-BIH, CPSC2019, <i>Europen ST-T</i> e INCART não há 12 derivações.	x		x	x	x	x	x	x			
Rotulagem da Base de dados	Rotulagem baseada na marcação dos picos R-R		x		x	x	x	x	x	x		
	Rotulagem baseada na quantidade de pontos amostrados do complexo QRS			x								
	Rotulagem manual de início e término do complexo QRS										x	x
	Procedimento de rotulagem não especificado	x										
Pré-processamento	Operação de diferença e média	x										
	Filtros: Mediano e passa-baixa (<i>butterworth</i>)		x		x	x						
	Normalização de dados e reamostragem								x	x		x
	Utilização de filtros baseados em <i>wavelets</i> ou redes neurais					x	x	x				
	Não utiliza pré-processamentos especiais			x							x	
Arquitetura do Modelo	Arquiteturas CNN com imagem do sinal de ECG		x									
	Arquiteturas que envolvem CNN e outros modelos	x			x	x	x					
	Arquiteturas com o paradigma temporal em sequências de eventos.	x			x	x	x					
	Modelos que aprendem características em diferentes resoluções espaciais	x			x			x	x	x	x	x

	Aplica validação cruzada			x									
	Intervalo do segmento \overline{QS} não definido	x			x	x	x						
	Intervalo do segmento \overline{QS} definido							x	x	x	x	x	
Pós-processamento	Não faz uso de pós-processamento	x	x									x	
	Utilização de modelos de redes neurais para o pós-processamento				x		x		x				
	Utilização de limiar específico aplicado à saída da rede							x		x			
	Utilização do limiar baseado no pico R-R					x			x				x
	Resposta final do tipo binária no intervalo [0-1]				x	x		x	x	x	x	x	x

6 CONCLUSÕES

Este estudo avançou no estado da arte na detecção do complexo QRS nas 12 derivações do ECG, utilizando redes neurais convolutivas de segmentação semântica, associadas a técnicas de pós-processamento. Foi realizado um estudo de ablação, com três arquiteturas de redes convolutivas, uma com 25 filtros, outra com 32 filtros e uma terceira com 75 filtros nas camadas convolutivas de extração de características. O melhor desempenho foi obtido com a arquitetura de 32 filtros nas camadas convolutivas. O método de pós-processamento adotado nesta pesquisa elimina intervalos com formas de onda cuja amplitude é menor do que a metade da média das amplitudes dos complexos QRS de uma dada amostra. Com essa arquitetura de rede convolutiva associada ao pós-processamento, os melhores resultados médios obtidos para as métricas, precisão, *recall* e *F1-score*, nas 12 derivações foram 99,98%, 99,90% e 99,94%, respectivamente.

O principal diferencial desta pesquisa foi permitir avaliar a duração do intervalo \overline{QS} , que é importante no diagnóstico de doenças cardiovasculares. Verificou-se que, mesmo em intervalos ruidosos, o modelo consegue obter com precisão a duração do complexo QRS. No entanto, em amostras que apresentam infarto do miocárdio precoce, o complexo QRS é mais largo. Nesses casos, o modelo detecta um comprimento menor para o intervalo \overline{QS} . Acredita-se que, acrescentando mais amostras de infarto do miocárdio no treinamento, consigamos resolver esse problema.

6.1 Trabalhos Futuros

A experiência obtida com o trabalho desenvolvido, encaminhou-nos naturalmente para a proposição de trabalhos futuros que visam o aperfeiçoamento do trabalho ora desenvolvido. Seguem as principais sugestões de trabalhos futuros:

- Utilizar mais dados que contenham os intervalos característicos de infarto do miocárdio precoce com complexos QRS mais largos;
- Utilizar os pesos aprendidas com as derivações individuais para a construção de um único modelo aplicável a todas as derivações;

- Treinar um modelo que tenha como entrada dados de todas as 12 derivações;
- Aplicar o modelo para a detecção de segmentos de ondas P e T nas 12 derivações;
- Obter a amplitude e a largura do complexo QRS, com vistas ao cálculo de parâmetros para o diagnóstico de doenças cardiovasculares;
- Aperfeiçoar a detecção do intervalo \overline{QS} em amostras com infarto do miocárdio.

7 REFERÊNCIAS

BELKADI, Mohamed Amine; DAAMOUCHE, Abdelhamid; MELGANI, Farid. A deep neural network approach to QRS detection using autoencoders. *Expert Systems with Applications*, v. 184, p. 115528, 2021.

BORDE, Anna; KOLOKOLNIKOV, George; SKURATOV, Victor. Development of neural network-based approach for QRS segmentation. In: 2019 25th Conference of Open Innovations Association (FRUCT). IEEE, 2019. p. 77-84.

CAI, Wenjie; HU, Danqin. QRS complex detection using novel deep learning neural networks. *IEEE Access*, v. 8, p. 97082-97089, 2020.

CHARU, C. Aggarwal. *Neural networks and deep learning: a textbook*. 2018.

DALVI, R. F. Suporte a detecção e classificação de câncer a partir de mamografias digitalizadas e redes neurais convolucionais. *Dissertação (Mestrado em Engenharia Elétrica) - Centro Tecnológico, Universidade Federal do Espírito Santo*. Vitória, p. 106. 2018.

DE OLIVEIRA, F. Ivan; CORTEZ, P. Usar. A QRS detection based on Hilbert transform and wavelet bases. In: *Proceedings of the 2004 14th IEEE Signal Processing Society Workshop Machine Learning for Signal Processing*, 2004. IEEE, 2004. p. 481-489.

DURAJ, Konrad *et al.* Semantic Segmentation of 12-Lead ECG Using 1D Residual U-Net with Squeeze-Excitation Blocks. *Applied Sciences*, v. 12, n. 7, p. 3332, 2022.

EKMAN, Magnus. *Learning Deep Learning: Theory and Practice of Neural Networks, Computer Vision, NLP, and Transformers Using TensorFlow*. Addison-Wesley Professional, 2021.

FOTOOHINASAB, Atiyeh; HOCKING, Toby; AFGHAH, Fatemeh. A greedy graph search algorithm based on changepoint analysis for automatic QRS complex detection. *Computers in biology and medicine*, v. 130, p. 104208, 2021.

GÉRON, Aurélien. *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow*. " O'Reilly Media, Inc.", 2022.

GOLDBERGER, Ary. *Goldberger's Clinical Electrocardiography*. Philadelphia: Elsevier, 2018.

GOLDBERGER, Ary L. *et al.* PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *circulation*, v. 101, n. 23, p. e215-e220, 2000.

GONZALES, R. C.; WOODS, R. E. *Digital Image Processing*. Pearson. Harlow, England, 2018.

GOOVAERTS, Griet *et al.* A machine-learning approach for detection and quantification of QRS fragmentation. *IEEE journal of biomedical and health informatics*, v. 23, n. 5, p. 1980-1989, 2018.

HABIB, Ahsan; KARMAKAR, Chandan; YEARWOOD, John. Learning post-processing for QRS detection using Recurrent Neural Network. *arXiv preprint arXiv:2110.04130*, 2021.

HE, Runnan *et al.* Automatic detection of QRS complexes using dual channels based on U-Net and bidirectional long short-term memory. *IEEE Journal of Biomedical and Health Informatics*, v. 25, n. 4, p. 1052-1061, 2020.

HE, Zhengling *et al.* A Novel Convolutional Neural Network for Arrhythmia Detection From 12-lead Electrocardiograms. In: *2020 Computing in Cardiology*. IEEE, 2020. p. 1-4.

HU, Xiao *et al.* Detection of onset and offset of QRS complex based a modified triangle morphology. In: *Frontier and Future Development of Information Technology in Medicine and Education*. Springer, Dordrecht, 2014. p. 2893-2901.

HUANG, Boqiang; WANG, Yuanyuan. QRS complexes detection by using the principal component analysis and the combined wavelet entropy for 12-lead electrocardiogram signals. In: *2009 Ninth IEEE international conference on computer and information technology*. IEEE, 2009. p. 246-251.

LI, Cuiwei; ZHENG, Chongxun; TAI, Changfeng. Detection of ECG characteristic points using wavelet transforms. *IEEE Transactions on biomedical Engineering*, v. 42, n. 1, p. 21-28, 1995.

JALIL, Bushra *et al.* QRS Complex Detection by Non-Linear Thresholding of Modulus Maxima. In: *2010 20th International Conference on Pattern Recognition*. IEEE, 2010. p. 4500-4503.

JIA, Menghan *et al.* Robust QRS detection using high-resolution wavelet packet decomposition and time-attention convolutional neural network. *IEEE Access*, v. 8, p. 16979-16988, 2020.

KINGMA, D.; BA, J. Adam: A Method for Stochastic Optimization. International Conference on Learning Representations, 2014.

KOSCOVA, Zuzana *et al.* QRS Complex Detection in Paced and Spontaneous Ultra-High-Frequency ECG. In: 2021 Computing in Cardiology (CinC). IEEE, 2021. p. 1-4.

LONG, J.; SHELHAMER, E.; DARRELL, T. Fully convolutional networks for semantic segmentation. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, v. 07–12–June, p. 3431–3440, 2015.

LIU, Wei *et al.* An Octave Convolution Neural Network-based QRS Detector. In: 2020 International Conference on Sensing, Measurement & Data Analytics in the era of Artificial Intelligence (ICSMD). IEEE, 2020. p. 413-418.

LU, Xuanyu; PAN, Maolin; YU, Yang. QRS detection based on improved adaptive threshold. Journal of healthcare engineering, v. 2018, 2018.

LU, Xuanyu; PAN, Maolin; YU, Yang. QRS detection based on improved adaptive threshold. Journal of healthcare engineering, v. 2018, 2018.

LUTFULLAH, Zubair *et al.* 12 lead QRS window detection: Using feature extraction and statistical parameters. In: 2011 5th International Conference on Bioinformatics and Biomedical Engineering. IEEE, 2011. p. 1-4.

MEHTA, S. S.; LINGAYAT, N. S. SVM based QRS detection in electrocardiogram using signal entropy. IETE Journal of Research, v. 54, n. 3, p. 231-240, 2008.

MOFFA, P.; SANCHES, P. Eletrocardiograma: Uma abordagem didática. 1ª edição. Rocca, São Paulo, p. 356, 2010.

MUSA, Nehemiah *et al.* A systematic review and Meta-data analysis on the applications of Deep Learning in Electrocardiogram. Journal of ambient intelligence and humanized computing, p. 1-74, 2022.

NAGAL, Devendra; SHARMA, Swati. Simultaneous 12-lead QRS detection by K-means clustering algorithm. In: International Conference on Recent Advances and Innovations in Engineering (ICRAIE-2014). IEEE, 2014. p. 1-4.

PAN, Jiapu; TOMPKINS, Willis J. A real-time QRS detection algorithm. IEEE transactions on biomedical engineering, n. 3, p. 230-236, 1985.

PANDIT, Diptangshu *et al.* A lightweight QRS detector for single lead ECG signals using a max-min difference algorithm. *Computer methods and programs in biomedicine*, v. 144, p. 61-75, 2017.

PARALIC, Martin. Detection of QRS complexes using convolutional neural network. In: 2019 42nd International Conference on Telecommunications and Signal Processing (TSP). IEEE, 2019. p. 182-186.

SILVA, Pedro *et al.* QRS detection in ECG signal with convolutional network. In: Iberoamerican Congress on Pattern Recognition. Springer, Cham, 2018. p. 802-809.

Timmis, Adam, *et al.*, "European Society of Cardiology: cardiovascular disease statistics 2019.", *European heart journal*, 41, 12-85 (2020).

WANG, Xingjun; ZOU, Qingyan. QRS detection in ECG signal based on residual network. In: 2019 IEEE 11th International Conference on Communication Software and Networks (ICCSN). IEEE, 2019. p. 73-77.

World Health Organization, " Dialogue with the private sector on medicines and health technologies for hypertension, June 2022",WHO, 11 June 2021, < <https://www.who.int/news-room/events/detail/2022/06/21/default-calendar/dialogue-with-the-private-sector-on-medicines-and-health-technologies-for-hypertension--june-2022> > (22 fevereiro 2023).

XIANG, Yande; LIN, Zhitao; MENG, Jianyi. Automatic QRS complex detection using two-level convolutional neural network. *Biomedical engineering online*, v. 17, n. 1, p. 1-17, 2018.

YU, Runze *et al.* QRS detection and measurement method of ECG paper based on convolutional neural networks. In: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2018. p. 4636-4639

YUEN, Brosnan; DONG, Xiaodai; LU, Tao. Inter-patient CNN-LSTM for QRS complex detection in noisy ECG signals. *IEEE Access*, v. 7, p. 169359-169370, 2019.

ZHANG, Aston *et al.* Dive into deep learning. arXiv preprint arXiv:2106.11342, 2021.

ZHAO, Wei *et al.* A multilead fusion based QRS complex detection method on 12-lead electrocardiogram signals. In: 2018 Computing in Cardiology Conference (CinC). IEEE, 2018. p. 1-4.

ZHONG, Wei *et al.* A deep learning approach for fetal QRS complex detection. *Physiological measurement*, v. 39, n. 4, p. 045004, 2018.

ZHOU, Zhi-Hua. *Machine learning*. Springer Nature, 2021.