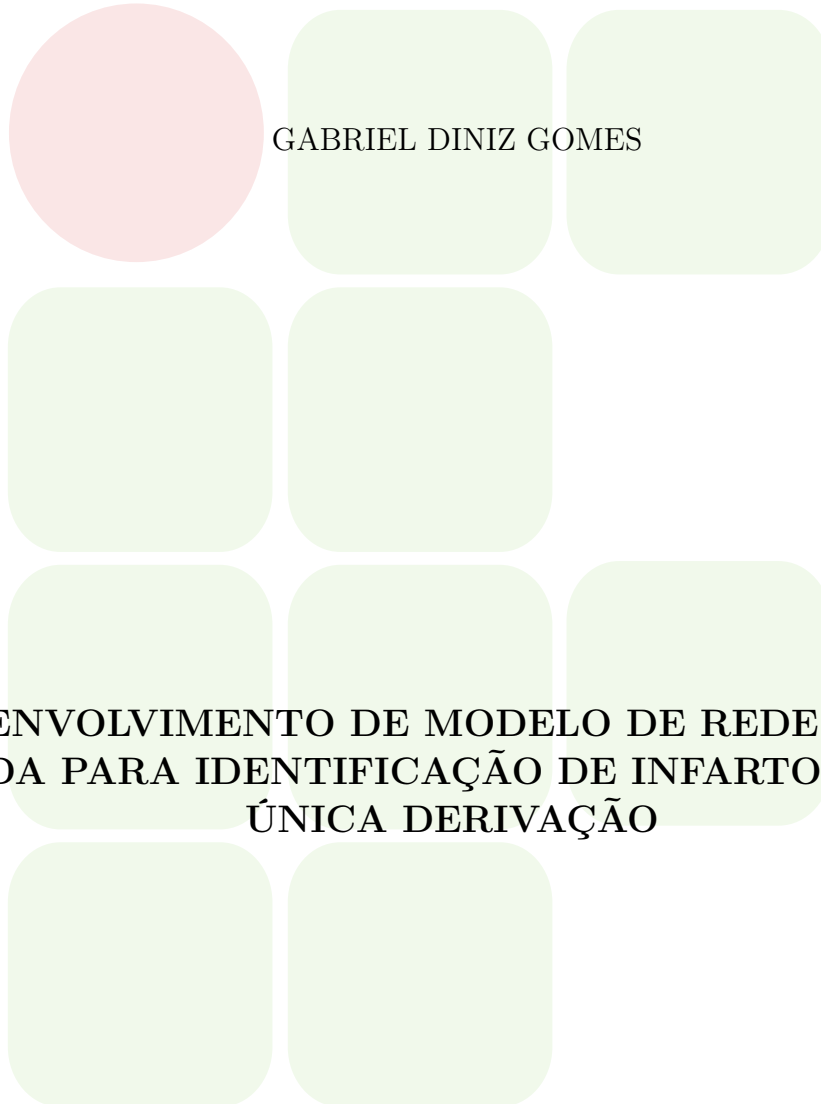


INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DA
PARAÍBA

COORDENAÇÃO DO CURSO DE ENGENHARIA ELÉTRICA



GABRIEL DINIZ GOMES

**DESENVOLVIMENTO DE MODELO DE REDE NEURAL
HÍBRIDA PARA IDENTIFICAÇÃO DE INFARTO COM UMA
ÚNICA DERIVAÇÃO**

João Pessoa

2023

GABRIEL DINIZ GOMES

**DESENVOLVIMENTO DE MODELO DE REDE NEURAL
HÍBRIDA PARA IDENTIFICAÇÃO DE INFARTO COM UMA
ÚNICA DERIVAÇÃO**

Trabalho de Conclusão de Curso submetido à Coordenação do Curso de Engenharia Elétrica do Instituto Federal de Educação, Ciência e Tecnologia da Paraíba, como parte dos requisitos para a obtenção do grau de Engenheiro Eletricista.

Orientador: Carlos Danilo Miranda Regis

João Pessoa
2023

Dados Internacionais de Catalogação na Publicação – CIP
Biblioteca Nilo Peçanha – IFPB, *Campus* João Pessoa

G633d	Gomes, Gabriel Diniz Desenvolvimento de modelo de rede neural híbrida para identificação de infarto com uma única derivação / Gabriel Diniz Gomes. – 2023. 39 f. TCC (Graduação – Bacharelado em Engenharia Elétrica) – Instituto Federal da Paraíba – IFPB / Coordenação de Engenharia Elétrica, 2023. Orientador: Prof ^o Dr. Carlos Danilo Miranda Regis. 1. Rede neural. 2. Eletrocardiograma. 3. Infarto do miocárdio. 4. Rede convolucional. 5. Rede Recorrente. I. Título. CDU 004.032.26:616.12-073.7
-------	---

Bibliotecária responsável Ivanise Andrade Melo de Almeida – CRB15/96

GABRIEL DINIZ GOMES

DESENVOLVIMENTO DE MODELO DE REDE NEURAL HÍBRIDA PARA IDENTIFICAÇÃO DE INFARTO COM UMA ÚNICA DERIVAÇÃO

Trabalho de Conclusão de Curso submetido à Coordenação do Curso de Engenharia Elétrica do Instituto Federal de Educação, Ciência e Tecnologia da Paraíba, como parte dos requisitos para a obtenção do grau de Engenheiro Eletricista.

BANCA EXAMINADORA

Documento assinado digitalmente



Carlos Danilo Miranda Regis
Data: 21/12/2023 10:52:29-0300
Verifique em <https://validar.iti.gov.br>

Carlos Danilo Miranda Regis, Dr. – IFPB
Orientador

Documento assinado digitalmente



PATRIC LACOUTH DA SILVA
Data: 21/12/2023 14:16:43-0300
Verifique em <https://validar.iti.gov.br>

Patric Lacouth Da Silva, Dr. – IFPB
Membro da Banca

Documento assinado digitalmente



Suzete Elida Nobrega Correia
Data: 23/12/2023 07:24:47-0300
Verifique em <https://validar.iti.gov.br>

Suzete Élide Nóbrega Correia, Dra. – IFPB
Membro da Banca

João Pessoa, 12 de dezembro de 2023.

Agradecimentos

Agradeço primeiramente a Deus. Sem Ele a caminhada ao longo curso, o aprendizado de cada disciplina e a superação de cada desafio jamais teriam sido possíveis.

Agradeço aos meus pais por todo o apoio durante essa trajetória de minha vida, sempre me incentivando nos estudos e me provendo a estrutura para continuar minha caminhada e seguir meus sonhos. Vocês são minha maior inspiração e exemplo. Espero um dia ser capaz de retribuir tudo que fizeram por mim.

À minha noiva, melhor amiga e amor da minha vida, Myllena Lys, agradeço por acompanhar-me desde o início desta jornada, sendo compreensiva, amorosa e companheira, me aconselhando nos momentos de necessidade e me suportando nos períodos de dificuldade. Você torna minha vida mais leve e prazerosa.

Ao meu orientador, o professor Danilo Regis, expresso minha sincera gratidão. Suas orientações e aconselhamentos sempre estiveram presentes ao longo das disciplinas e trabalhos realizados, contribuindo significativamente para meu crescimento acadêmico e profissional.

RESUMO

Doenças cardiovasculares são a maior causa de mortes no mundo, sendo responsáveis por cerca de 17,9 milhões de mortes anualmente. O infarto agudo do miocárdio (IAM) é caracterizado pela oclusão de uma artéria coronária, a qual pode levar à necrose do músculo cardíaco pela falta de oxigenação decorrente da ausência de fluxo sanguíneo. O principal exame utilizado no diagnóstico do IAM é o eletrocardiograma (ECG), a partir do qual diferentes perspectivas da atividade elétrica do coração podem ser monitoradas. O presente trabalho teve por objetivo o desenvolvimento de uma rede neural híbrida, empregando redes neurais convolucionais e recorrentes, capaz de identificar o IAM em sinais de ECG utilizando uma única derivação eletrocardiográfica, visando a implementação em dispositivos vestíveis, como *smartwatches*. Para isso, buscou-se identificar qual comprimento de sinal e conjunto de hiperparâmetros apresentaria as melhores métricas de desempenho ao receber os sinais de ECG como entrada da rede. Ao aplicar o conjunto de teste utilizando o comprimento de sinal e conjunto de hiperparâmetros ideais, observou-se uma acurácia de 84,23%, precisão de 74,47%, sensibilidade de 74,19% e F1-Score de 74,33%.

Palavras-chave: Infarto do Miocárdio. Eletrocardiograma. Derivação Única. Rede Convolutacional. Rede Recorrente.

ABSTRACT

Cardiovascular diseases are the leading cause of death worldwide, accounting for approximately 17.9 million deaths annually. Acute myocardial infarction (AMI) is characterized by the occlusion of a coronary artery, leading to myocardial necrosis due to oxygen deprivation caused by the absence of blood flow. The primary diagnostic tool for AMI is the electrocardiogram (ECG), which allows monitoring different perspectives of the heart's electrical activity. This study aimed to develop a hybrid neural network, employing both convolutional and recurrent neural networks, capable of identifying AMI in ECG signals using a single electrocardiographic lead, with the goal of implementation in wearable devices such as smartwatches. To achieve this, the optimal signal length and hyperparameter set that exhibited the best performance metrics when receiving ECG signals as input to the network were identified. Applying the test dataset with the ideal signal length and hyperparameter set resulted in an accuracy of 84.23%, precision of 74.47%, sensitivity of 74.19%, and F1-Score of 74.33%.

Keywords: *Myocardial Infarction. Electrocardiogram. Single Lead. Convolutional Network. Recurrent Network.*

LISTA DE FIGURAS

Figura 1 – Representação da Doença Arterial Coronariana.	12
Figura 2 – ECG de 12 derivações.	13
Figura 3 – Posições dos vetores de derivações no corpo humano. (À esquerda) Vetores das derivações periféricas. (À direita) Vetores das derivações precordiais.	13
Figura 4 – Eletrocardiograma e seus componentes característicos.	14
Figura 5 – Sinal de ECG com Infarto Agudo do Miocárdio.	15
Figura 6 – Exemplo do processo de convolução unidimensional para uma entrada (<i>Sample</i>) composta por 13 amostras, com o filtro da média móvel (<i>kernel</i>) de tamanho 7 e saída (<i>Output</i>) com a dimensionalidade reduzida para 7 amostras.	17
Figura 7 – Extração de padrões e redução de dimensionalidade por meio das camadas convolucionais e de subamostragem.	18
Figura 8 – Arquitetura de uma RNN. (À esquerda) A RNN e o cálculo de sua função de perda desenhadas com conexões recorrentes. (À direita) O mesmo visto como um grafo computacional desdobrado no tempo, onde cada nó está agora associado a uma instância de tempo específica. . . .	19
Figura 9 – Arquitetura de uma LSTM.	20
Figura 10 – Diagrama de funcionamento do trabalho.	22
Figura 11 – Divisão dos sinais do banco em suas superclasses e subclasses.	23
Figura 12 – Arquitetura da rede neural elaborada.	26
Figura 13 – Função ReLU e função sigmoide.	27
Figura 14 – Amostra de sinal de ECG de 3 s.	28
Figura 15 – Evolução da função de perda (<i>Loss</i>) e AUC durante o treino do modelo final da rede.	32
Figura 16 – Matriz de confusão para classificação de ECG.	33

LISTA DE TABELAS

Tabela 1 – Hiperparâmetros ótimos para os diferentes comprimentos de sinais de ECG. (<i>Kernel</i> diz respeito ao tamanho do filtro de convolução, LSTM corresponde ao tamanho da saída da camada correspondente e LR diz respeito à taxa de aprendizado.)	30
Tabela 2 – Resultados das métricas para os diferentes comprimentos de sinais de ECG.	30
Tabela 3 – Resultados dos desvios padrão para os diferentes comprimentos de sinais de ECG.	31
Tabela 4 – Resultado do teste da rede.	32

SUMÁRIO

1	INTRODUÇÃO	9
2	INFARTO AGUDO DO MIOCÁRDIO	11
2.1	Eletrocardiograma	12
2.2	Impactos do IAM no sinal de ECG	14
3	REDES NEURAIS	16
3.1	Redes Neurais Convolucionais	16
3.2	Redes Neurais Recorrentes	18
3.2.1	Memória Longa de Curto Prazo	19
4	METODOLOGIA	22
4.1	Aquisição dos sinais	23
4.2	Processamento dos sinais	24
4.3	Configuração da rede neural	25
4.3.1	Seleção de hiperparâmetros da rede neural	27
4.4	Métricas de desempenho	28
5	RESULTADOS	30
5.1	Teste da rede	31
5.2	Discussões	33
6	CONCLUSÃO	35
	REFERÊNCIAS	36

1 Introdução

De acordo com a Organização Mundial da Saúde (OMS), as doenças cardiovasculares representam a principal causa de óbitos globalmente. Em 2019, aproximadamente 17,9 milhões de pessoas perderam a vida devido a doenças arteriais coronarianas, dentre as quais 85% dessas fatalidades foram atribuídas ao infarto ou acidente vascular cerebral (ORGANIZATION, 2017; BRAUNWALD et al., 2019). Apesar da significativa taxa de mortalidade associada a essas condições, destaca-se a possibilidade de prevenção eficaz mediante o avanço das técnicas de diagnóstico e tratamento (MOZAFFARIAN et al., 2016).

O Eletrocardiograma (ECG) é o exame mais utilizado para identificar e diagnosticar cardiopatias, por ser rápido, barato e não invasivo (FELDMAN; GOLDWASSER, 2004). Para a obtenção do ECG, são dispostos eletrodos sobre a pele do paciente, que registram as variações de potencial elétrico do coração a partir de diferentes perspectivas em um mesmo intervalo de tempo. Essas diferentes perspectivas são denominadas derivações, e há 12 delas, que permitem que a atividade elétrica do coração seja analisada, possibilitando identificar um infarto agudo do miocárdio (IAM), arritmias e outras patologias que afetem a condução elétrica. Essa análise, entretanto, é feita para mais de um batimento cardíaco e de forma qualitativa (HAGIWARA et al., 2018), sendo portanto, sujeita a erros por parte do examinador, por ser dependente de sua experiência. Em grande parte dos casos deve ser realizada de forma rápida, como o caso de um IAM, que requer intervenção imediata e pode causar danos irreversíveis ao músculo cardíaco, caso dure mais de duas horas (THYGESEN et al., 2007; GAZIANO et al., 2006).

Redes neurais artificiais e técnicas de aprendizado de máquina têm desempenhado um papel fundamental na busca por aprimorar a detecção e diagnóstico de condições e patologias cardiovasculares. Métodos baseados em algoritmos capazes de aprender padrões complexos a partir de dados têm se destacado na interpretação de sinais biomédicos, como os provenientes de ECG. Trabalhos notáveis nesse sentido podem ser encontrados em Joloudari et al. (2022) e Petmezas et al. (2022), onde observa-se que redes neurais convolucionais (CNNs) e recorrentes (RNNs), em particular, têm mostrado eficácia em reconhecer padrões sutis nos sinais cardíacos, contribuindo para uma análise mais precisa e automatizada. A capacidade de processamento dessas tecnologias oferece potencial para avançar na identificação precoce de problemas cardíacos, permitindo intervenções rápidas e eficazes.

A evolução tecnológica tem possibilitado o desenvolvimento de dispositivos de monitoramento de saúde cada vez mais acessíveis e integrados ao cotidiano, como *smartwatches*, que têm se tornado ferramentas promissoras para o monitoramento contínuo da

saúde cardiovascular. A incorporação de sistemas de diagnóstico de ECG em *smartwatches* oferece a oportunidade de monitorar constantemente a atividade cardíaca, proporcionando aos usuários informações em tempo real sobre sua saúde cardiovascular. Além da praticidade, essa abordagem tem o potencial de detectar anomalias precoces e alertar os usuários sobre possíveis problemas, permitindo uma resposta rápida e reduzindo os riscos associados a eventos cardíacos.

Trabalhos na área de classificação de sinais de ECG têm obtido ótimas métricas fazendo uso de mais de uma derivação eletrocardiográfica, como o desenvolvido por He et al. (2021), onde foi implementada uma rede neural para identificação de infarto utilizando as 12 derivações do sinal de ECG, onde cada derivação é tratada por uma rede individual, especializada na extração de características de sua derivação. Com isso, He et al. (2021) foram capazes de alcançar uma acurácia de 94,19%, além de concluir que o sistema pode realizar o diagnóstico em tempo real utilizando a plataforma ARM Cortex-A72.

Já o trabalho desenvolvido por Vieira et al. (2022) utilizou uma rede MLP para identificar infarto, fazendo uso de técnicas de pré-processamento que envolvem o uso das 12 derivações eletrocardiográficas para possibilitar a obtenção do vetocardiograma e do espaço de fases, passando, portanto, por uma etapa extensiva de pré-processamento, a fim de posteriormente obter as características dos sinais através de descritores de Haralick para que estas atuem como entrada da rede. O melhor resultado desta rede obteve as métricas de acurácia de 70,31% e sensibilidade de 69,76%.

O presente trabalho teve por objetivo o desenvolvimento de uma rede neural capaz de identificar o IAM em sinais de ECG fazendo o uso de apenas uma derivação, visando a possibilidade futura de implementação do sistema em dispositivos vestíveis, como *smartwatches*. Para isso, foi desenvolvida uma rede neural híbrida, usando CNN e RNN para a extração de características dos sinais cardíacos.

O Capítulo 2 descreve as características e epidemiologia do infarto agudo do miocárdio. Em seguida, no seção 2.1, será explorado o funcionamento do eletrocardiograma, bem como sua aquisição e o impacto do IAM sobre ele. O Capítulo 3 detalhará o funcionamento de uma Rede Neural Convolutiva e de uma Rede Neural Recorrente, bem como os conceitos que envolvem a sua operação. No Capítulo 4 serão descritos o conjunto de dados utilizado, o processamento ao qual os sinais foram submetidos, a configuração da rede neural e a sua seleção de hiperparâmetros e, por fim, as métricas utilizadas para a avaliação do seu desempenho. No Capítulo 5 serão expostos os resultados e, por fim, as conclusões obtidas e possíveis trabalhos futuros serão apresentadas no Capítulo 6.

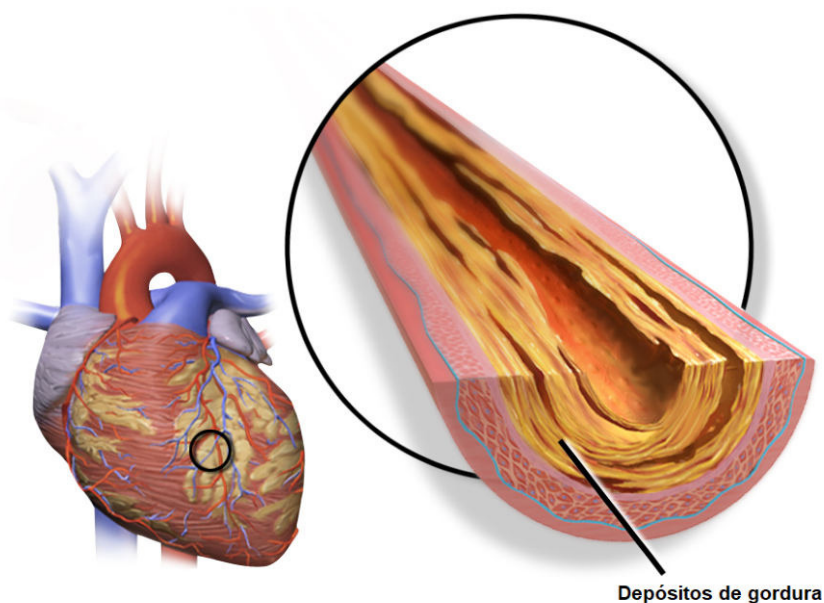
2 Infarto Agudo do Miocárdio

As doenças cardiovasculares, também chamadas cardiopatias, englobam um conjunto de condições que afetam o músculo cardíaco (miocárdio) e o sistema vascular, incluindo as artérias coronárias que fornecem sangue ao coração e ao cérebro (GAZIANO et al., 2006). Entre essas condições, destaca-se o infarto agudo do miocárdio (IAM), devido à sua elevada taxa de mortalidade nos primeiros 30 dias após o evento, afetando aproximadamente 33% dos pacientes, com metade dessas mortes ocorrendo antes que o paciente chegue a uma unidade de saúde (GAZIANO et al., 2006).

Durante o século XX, 80% das mortes por cardiopatias ocorreram em países subdesenvolvidos ou emergentes, e entre 1990 e 2013 elas aumentaram de 26% para 32% das mortes ao redor do mundo. Até a metade do século XX elas predominaram em países desenvolvidos, entretanto observa-se uma tendência de aumento em países de média e baixa renda, como o Brasil (BRAUNWALD et al., 2019; WHITE; CHEW, 2008). Tal transição epidemiológica ocorre devido ao processo de urbanização e à mudança no estilo de vida e trabalho nesses países, que causou um aumento nas taxas de obesidade, diabetes e hipertensão, fatores de risco para doenças cardíacas (BRAUNWALD et al., 2019).

O IAM ocorre quando o suprimento de oxigênio para o miocárdio é insuficiente para a demanda, e é causado pela oclusão total ou parcial de uma artéria coronária, podendo levar à necrose do músculo cardíaco pela falta de oxigênio e nutrientes decorrentes da ausência de fluxo sanguíneo caso dure, em média, de duas a quatro horas (THYGESEN et al., 2007; GAZIANO et al., 2006). Com isso, a capacidade de contração do músculo cardíaco é prejudicada, reduzindo, assim, o volume de sangue ejetado a cada batimento, bem como as capacidades físicas do indivíduo. Cerca de 70% dos casos fatais de IAM ocorrem devido ao acúmulo de gordura no interior das artérias coronárias, condição chamada de Doença Arterial Coronariana (DAC), onde formam-se coágulos que bloqueiam o suprimento de sangue para o miocárdio (WHITE; CHEW, 2008; THYGESEN et al., 2007), como pode ser observado na Figura 1.

Assim como todas as cardiopatias, o IAM tem como principal sintoma a dor torácica. Todavia tal sintoma, de forma isolada, não pode ser tomado como diagnóstico preciso para o infarto, se fazendo necessário exames para identificação de possíveis patologias cardíacas, sendo o ECG o mais utilizado. Através dele pode-se notar alterações na condução elétrica do coração, que podem indicar a ocorrência de um infarto ou de outras síndromes cardíacas (SCHWARZ, 2009). Seus efeitos nos sinais elétricos cardíacos serão vistos no seção 2.1.

Figura 1 – Representação da Doença Arterial Coronariana.

Fonte: Adaptada de (BLAUS, 2014).

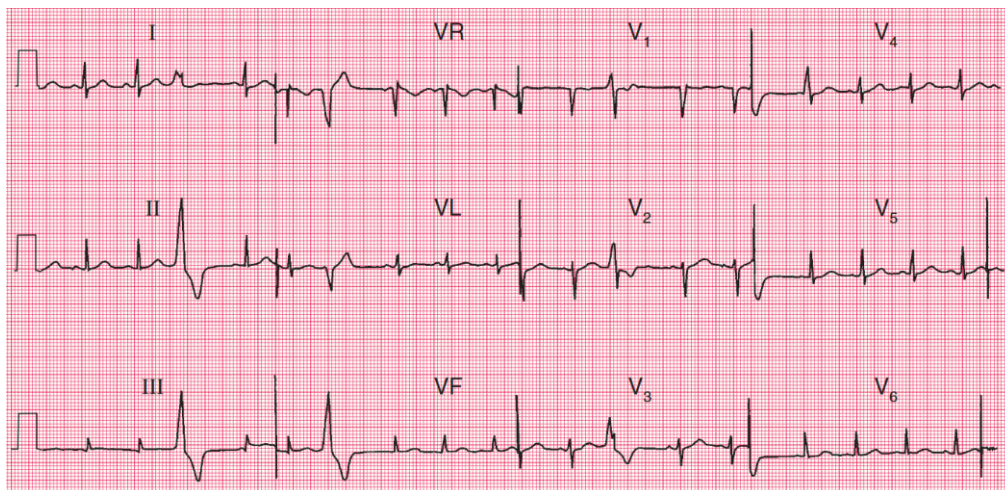
2.1 Eletrocardiograma

O eletrocardiograma (ECG) se trata de um exame cardíaco que mede os potenciais de campo elétrico gerados pelo coração e que estão relacionados às polarizações e despolarizações que ocorrem durante o ciclo cardíaco, sendo o resultado final de uma série complexa de processos fisiológicos e tecnológicos que envolvem a captação de correntes iônicas transmembrana geradas por fluxos de íons através das membranas celulares e entre células adjacentes. Essas correntes geram um campo elétrico variável no tempo, com significado fisiológico, dentro e ao redor do coração e que é modificado à medida que passa por diversas outras estruturas, como pulmões, sangue e músculo esquelético (BRAUNWALD et al., 2019). Para se obter o sinal de ECG, são posicionados eletrodos no corpo do paciente, que, por meio do eletrocardiógrafo, captam e amplificam os sinais para então registrá-los. Por ser um método não invasivo, de rápida aplicação e baixo custo, e de fácil transporte e manuseio, o ECG é um exame de rotina em clínicas para o diagnóstico de doenças cardíacas (CARDIOLOGIA, 2003).

Pode-se obter diferentes sinais de ECG de acordo com a disposição dos eletrodos sobre a pele do paciente. Essas disposições recebem o nome de derivações eletrocardiográficas e existem, basicamente, doze delas, sendo seis denominadas periféricas (DI, DII, DIII, aVR, aVL e aVF), pois representam o deslocamento vetorial das diferenças de potencial entre os eletrodos a partir do plano frontal do coração, captando o seu deslocamento para cima, para baixo, para a esquerda e direita. Já as outras seis derivações são chamadas precordiais (V1, V2, V3, V4, V5 e V6) e representam o deslocamento para frente e para trás, dando a noção de profundidade (REIS et al., 2013). A Figura 2 mostra essas 12

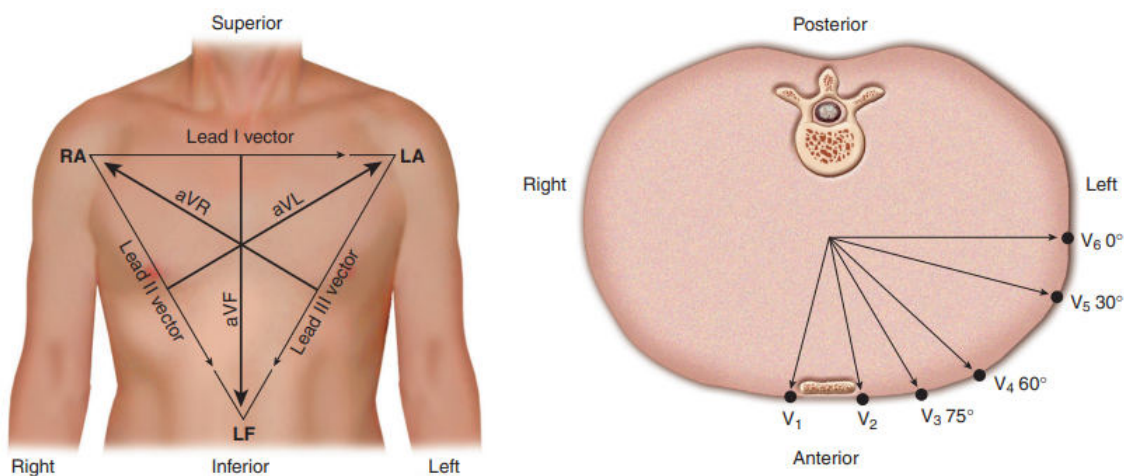
derivações, enquanto a Figura 3 ilustra os vetores de cada derivação no corpo humano, ou seja, a posição dos eletrodos no corpo para a captação de cada derivação.

Figura 2 – ECG de 12 derivações.



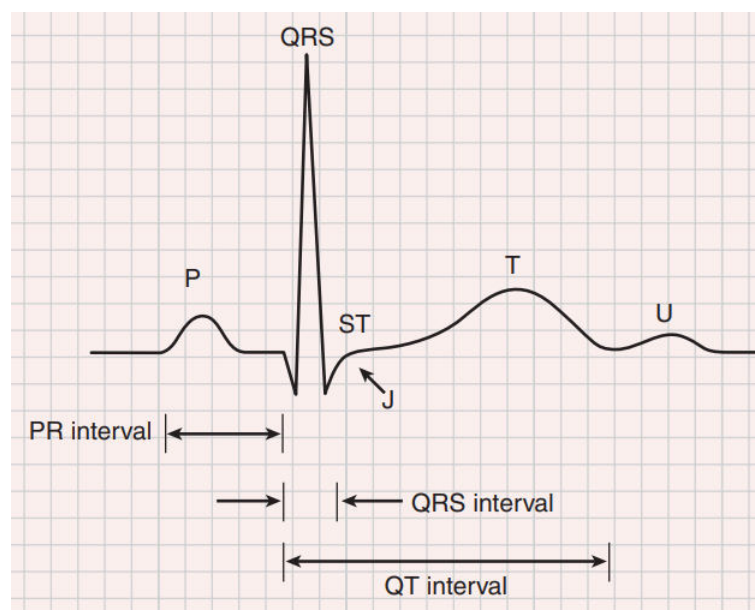
Fonte: (HAMPTON, 2014).

Figura 3 – Posições dos vetores de derivações no corpo humano. (À esquerda) Vetores das derivações periféricas. (À direita) Vetores das derivações precordiais.



Fonte: (BRAUNWALD et al., 2019).

Um sinal básico de ECG possui certos componentes que representam as fases que compõem o ciclo cardíaco, como observado na Figura 4, e que ocorrem em determinados pontos do coração. O ciclo se inicia com a onda P, que representa a despolarização dos átrios, seguido do intervalo PR, que diz respeito ao intervalo até que o estímulo elétrico chegue aos ventrículos. O complexo QRS representa a despolarização ventricular, que é seguido do segmento ST, que se trata do intervalo até o início da repolarização do ventrículo, que é finalmente denotada pela onda T (BRAUNWALD et al., 2019; PASTORE et al., 2016).

Figura 4 – Eletrocardiograma e seus componentes característicos.

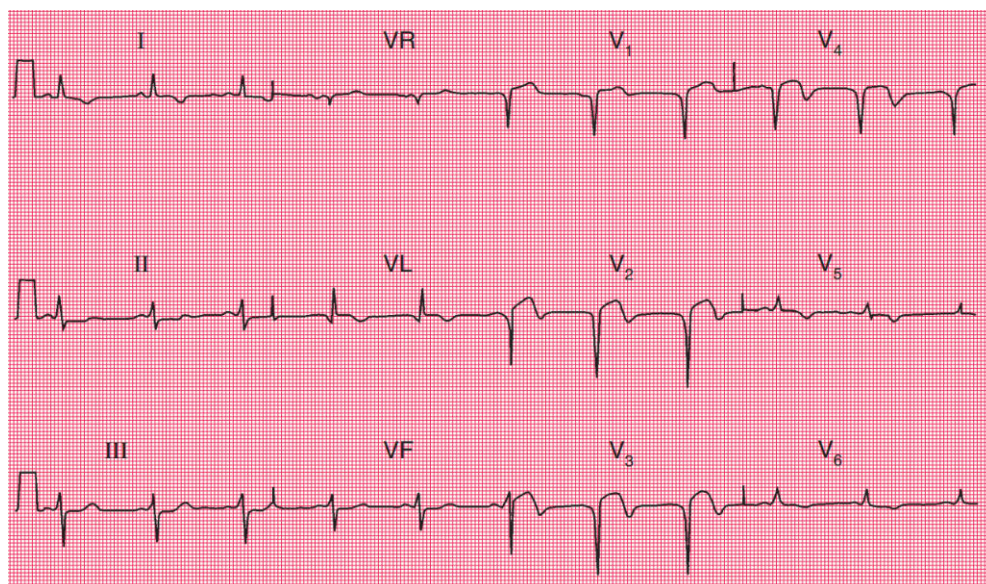
Fonte: (BRAUNWALD et al., 2019).

2.2 Impactos do IAM no sinal de ECG

Determinados padrões de ECGs, em conjunto com sintomas específicos, podem sugerir a presença de um IAM ou um infarto prévio. Tais anormalidades, que indicam a presença de IAM, podem ser observadas através do supradesnível no segmento ST (SST) em uma ou mais derivações contíguas, isto é, grupos de derivações de perspectivas adjacentes do coração e alterações na amplitude das ondas T maior ou igual a 0.1 mV, sendo este um sinal precoce que antecede a elevação do segmento ST, pois reflete o atraso de condução no miocárdio isquêmico (THYGESEN et al., 2007; PIEGAS et al., 2015).

Além disso, um infarto prévio pode ser observado através de ondas Q com tempo de duração superior a 0.02 s nas derivações V2 e V3 e ondas R com duração superior a 0.04 s nas derivações V1 e V2 (THYGESEN et al., 2007). Um exemplo de ECG com IAM pode ser observado na Figura 5, onde os sinais de infarto podem ser observados pelo supradesnivelamento do segmento ST nas derivações V2-V4.

Figura 5 – Sinal de ECG com Infarto Agudo do Miocárdio.



Fonte: (HAMPTON, 2014).

3 Redes Neurais

3.1 Redes Neurais Convolucionais

As redes neurais convolucionais (CNN, do inglês: *Convolutional Neural Network*) são uma classe especial de redes perceptron multicamadas (MLP, do inglês: *Multilayer Perceptron*) com foco no reconhecimento em duas dimensões de formas que possuem pouca variância à distorções como translação ou escalonamento (HAYKIN, 2009), sendo um tipo especializado de rede neural para processar dados que têm uma topologia conhecida, semelhante a uma matriz (GOODFELLOW; BENGIO; COURVILLE, 2016). Os exemplos mais comuns de utilização de CNNs são para o processamento de dados de imagem, como em (SILVA et al., 2023), onde as imagens podem ser consideradas uma matriz bidimensional de *pixels*, entretanto, dados de séries temporais também podem ser utilizados, sendo considerados uma matriz unidimensional de amostras em intervalos regulares de tempo (GOODFELLOW; BENGIO; COURVILLE, 2016).

O termo "rede neural convolucional" se dá por conta da utilização da operação de convolução pela rede. A convolução consiste em uma operação matemática fundamental usada em diversas áreas, como processamento de sinais, para filtragem de sinais, remoção de ruído ou realce de características; visão computacional, para detecção de bordas e reconhecimento de padrões; e aprendizado profundo, onde é a base das camadas convolucionais em CNNs (GOODFELLOW; BENGIO; COURVILLE, 2016). A convolução pode ser definida matematicamente a partir da Equação 3.1, sendo a integral do produto de f e g após g ser invertida e deslocada em relação a t . Aplicações computacionais, entretanto, utilizam dados discretizados no tempo, fazendo uso, portanto, da convolução linear discreta, cuja definição matemática pode ser observada na Equação 3.2 (HSU, 2004; OPPENHEIM; SCHAFER, 2013).

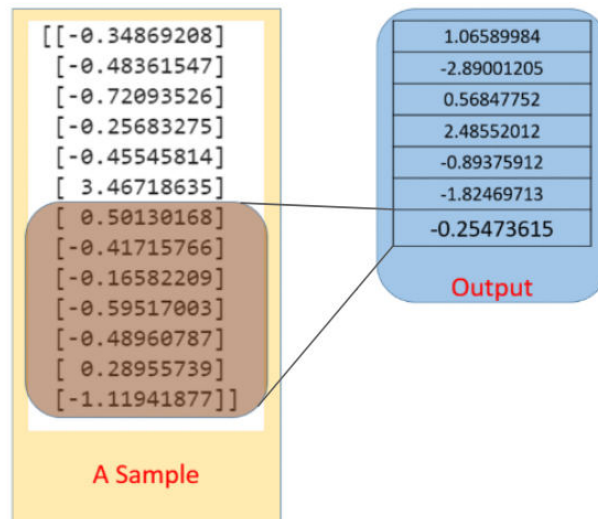
$$y(t) = (f * g)(t) = \int_{-\infty}^{\infty} f(\tau)g(t - \tau)d\tau \quad (3.1)$$

$$y[n] = (f * g)[n] = \sum_{m=-\infty}^{\infty} f[m] \cdot g[n - m] \quad (3.2)$$

Na terminologia das redes convolucionais, o primeiro argumento (função f) da convolução é frequentemente chamado de entrada, e o segundo argumento (função g) é chamado de filtro (ou *kernel*) e a saída (função y) é chamada de mapa de características (ou *feature map*) (GOODFELLOW; BENGIO; COURVILLE, 2016). Durante a convolução, o filtro desliza sobre os dados de entrada e realiza multiplicações elementares, somando os resultados para produzir um único valor em uma nova matriz, processo esse que pode ser

observado na Figura 6. Esse processo é repetido para todas as posições do filtro, criando assim o mapa de características. A operação de convolução permite que a rede aprenda automaticamente características importantes, como bordas e padrões, diretamente dos dados de entrada.

Figura 6 – Exemplo do processo de convolução unidimensional para uma entrada (*Sample*) composta por 13 amostras, com o filtro da média móvel (*kernel*) de tamanho 7 e saída (*Output*) com a dimensionalidade reduzida para 7 amostras.



Fonte: Adaptado de (KARAKAYA, 2021).

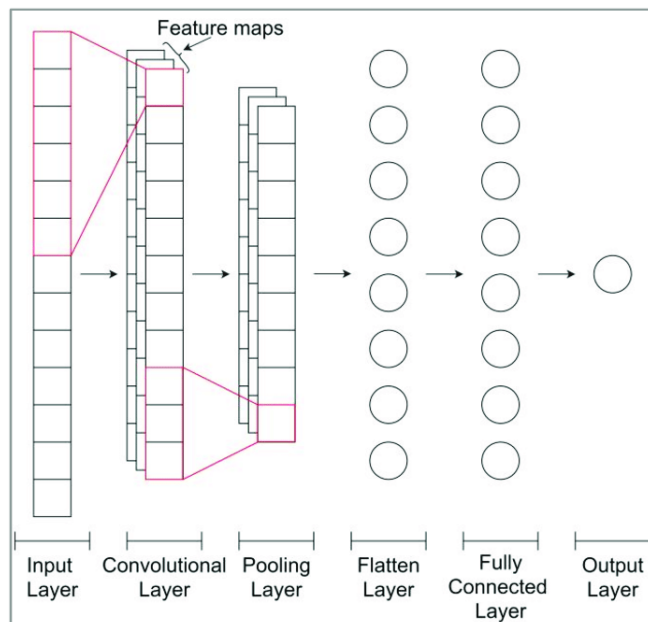
As CNNs possuem, portanto, a vantagem de serem capazes de reduzir a dimensionalidade do sinal de entrada ao mesmo tempo em que as suas características únicas para cada classe são capturadas. A extração de cada atributo é feita por meio de um neurônio que recebe como entradas apenas uma região da camada anterior, de forma que apenas características locais são extraídas e, assim, há uma redução da importância da posição delas no sinal, sendo mais importante a posição relativa dela em relação à outros atributos (HAYKIN, 2009; GOODFELLOW; BENGIO; COURVILLE, 2016). Outra vantagem das CNNs é o compartilhamento dos pesos sinápticos entre os neurônios, que leva a uma redução significativa no número de parâmetros, economizando recursos computacionais e de armazenamento e facilitando a generalização do modelo (HAYKIN, 2009).

Por fim, cada camada convolucional é, em geral, seguida de uma camada, chamada de camada de subamostragem (ou *pooling*) para a redução da resolução de seu resultado, de forma que um novo mapa de características com dimensões menores é obtido a partir da obtenção do valor máximo ou valor médio de cada mapa de características, reduzindo a sensibilidade dessa saída à distorções (HAYKIN, 2009).

Vários conjuntos dessas camadas convolucionais e de subamostragem podem ser adicionadas de forma encadeada, fazendo com que a complexidade dos atributos extraídos aumente à medida que essa cadeia aumenta, ao mesmo tempo em que o número de características e de pesos treináveis diferentes não aumenta expressivamente devido à

subamostragem e ao compartilhamento desses pesos. Esse processo de extração de características por meio de combinação dessas camadas pode ser observado na Figura 7.

Figura 7 – Extração de padrões e redução de dimensionalidade por meio das camadas convolucionais e de subamostragem.



Fonte: (NISA; KUAN, 2021).

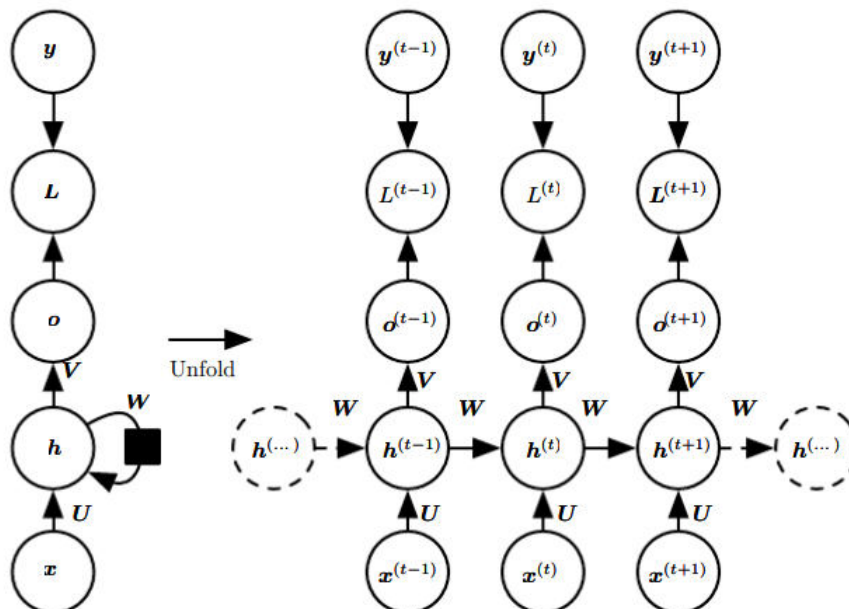
3.2 Redes Neurais Recorrentes

As redes neurais recorrentes (RNN, do inglês: *Recurrent neural networks*) constituem uma classe especial de arquiteturas de redes neurais profundas voltadas para o processamento de dados sequenciais. Diferente das CNNs, que podem facilmente se adaptar à imagens com grandes dimensões, as RNNs são dimensionadas para aplicações em sequências muito mais longas do que seria prático para redes sem esse tipo de especialização, como MLPs (GOODFELLOW; BENGIO; COURVILLE, 2016).

Uma RNN é composta por neurônios que mantêm uma memória interna que os permitem lidar com sequências, permitindo que a rede capture dependências temporais em dados. Cada neurônio em uma RNN recebe não apenas entradas do momento atual, mas também uma contribuição da memória anterior, tornando-as eficazes em tarefas que envolvem dados sequenciais, como séries temporais, processamento de áudio e processamento de sinais biológicos (HAYKIN, 2009; GOODFELLOW; BENGIO; COURVILLE, 2016). Sua arquitetura genérica pode ser visualizada na Figura 8, em que V , W e U são as matrizes de pesos, x é a sequência de entrada (*inputs*), o é a sequência de saída (*outputs*) e y é o alvo de treinamento desejado (*target*).

RNNs tradicionais, no entanto, têm limitações na retenção de informações de longo prazo devido ao problema do gradiente desaparecente ou explosivo. Esse problema está

Figura 8 – Arquitetura de uma RNN. (À esquerda) A RNN e o cálculo de sua função de perda desenhadas com conexões recorrentes. (À direita) O mesmo visto como um grafo computacional desdobrado no tempo, onde cada nó está agora associado a uma instância de tempo específica.



Fonte: (GOODFELLOW; BENGIO; COURVILLE, 2016).

diretamente relacionado aos pesos da rede, pois o gradiente é calculado durante o processo de retropropagação (*backpropagation*), onde o erro da rede é propagado para trás para ajustar os pesos, resultando em dois possíveis cenários problemáticos associados ao gradiente em RNNs: o gradiente desaparecente ou o gradiente explosivo (GOODFELLOW; BENGIO; COURVILLE, 2016). Ambos os problemas ocorrem em decorrência da natureza multiplicativa da atualização de pesos nas camadas temporais de uma RNN, de forma que, no caso do gradiente desaparecente, os gradientes diminuem exponencialmente à medida em que são multiplicados repetidamente pelos pesos da rede, resultando numa dificuldade de aprendizado da rede, pois suas atualizações nos pesos se tornam cada vez menores. Já no caso do gradiente explosivo, os gradientes aumentam exponencialmente ao longo da retropropagação, levando a atualizações de peso muito altas e tornando assim o treinamento instável, dificultando a convergência do modelo (HOCHREITER; BENGIO; FRASCONI, 2001; GOODFELLOW; BENGIO; COURVILLE, 2016).

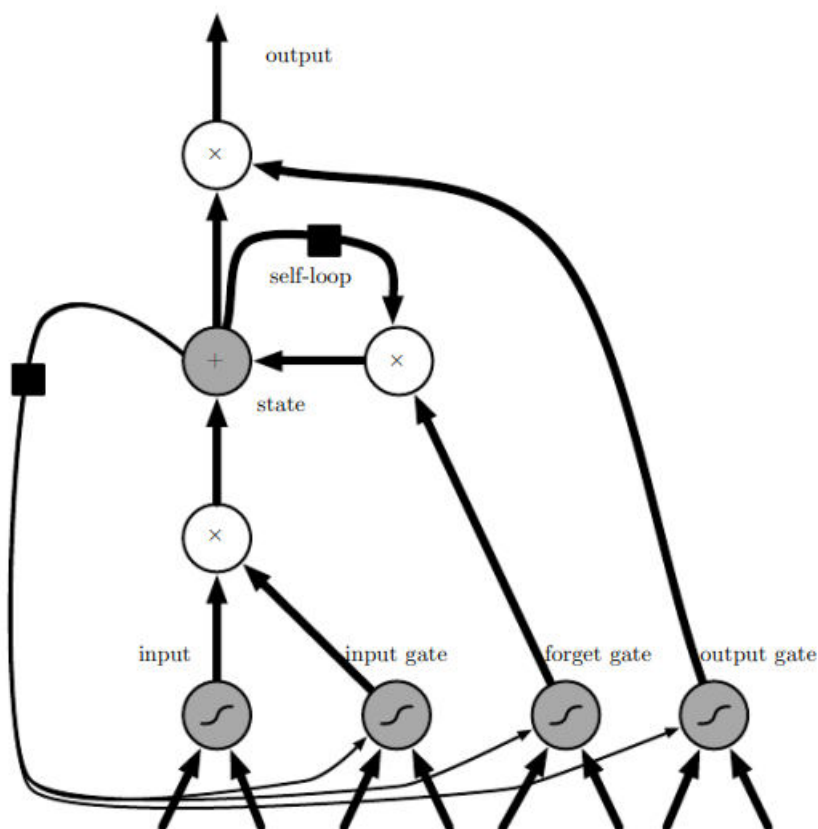
3.2.1 Memória Longa de Curto Prazo

A Memória Longa de Curto Prazo (LSTM, do inglês: *Long Short-Term Memory*) é um tipo de arquitetura de RNN projetada para lidar com o problema do gradiente desaparecente. Para isso, as LSTMs possuem uma arquitetura mais complexa, que as permitem capturar dependências temporais de longo prazo de maneira mais eficaz, permitindo lidar com sequências temporais mais longas (GOODFELLOW; BENGIO; COURVILLE, 2016).

Sua arquitetura consiste na introdução de "células LSTM"(ou "células de memória"), que têm uma recorrência interna (ou *auto-loop*) além da recorrência externa da RNN. Esses auto-loops, produzem caminhos nos quais o gradiente pode fluir por longas durações. Além disso, o peso desse *auto-loop* é adaptável ao invés de fixo, sendo este controlado por outra unidade oculta. Portanto, mesmo para uma LSTM com parâmetros fixos, a escala de tempo de integração, ou seja, o tempo o qual a célula lembrará de informações anteriores, pode mudar com base na sequência de entrada, porque as constantes de tempo são geradas pelo próprio modelo (HOCHREITER; BENGIO; FRASCONI, 2001; GOODFELLOW; BENGIO; COURVILLE, 2016).

Seu diagrama de blocos pode ser observado na Figura 9, em que cada célula tem as mesmas entradas e saídas de uma rede recorrente comum, entretanto com mais parâmetros e um sistema de unidades de controle que controla o fluxo de informações. Na LSTM, os pesos dos auto-loops são controlados por uma porta de esquecimento (ou *forget gate*), que define o peso para um valor entre 0 e 1 por meio de uma unidade sigmoide, evitando assim a explosão ou desaparecimento do gradiente.

Figura 9 – Arquitetura de uma LSTM.



Fonte: (GOODFELLOW; BENGIO; COURVILLE, 2016).

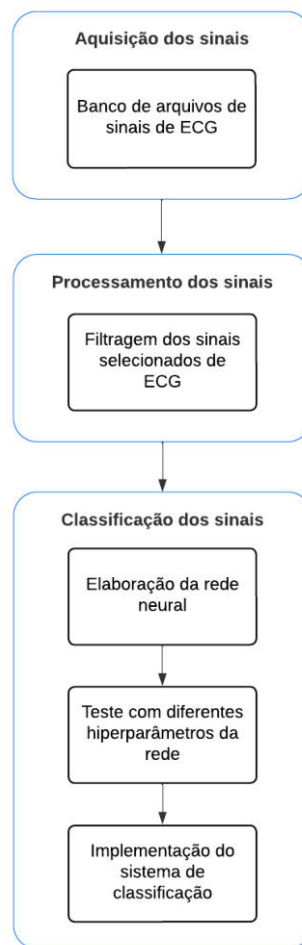
As LSTMs desempenham um papel crucial na modelagem de sequências temporais, tendo demonstrado aprender dependências de longo prazo com mais facilidade do que as arquiteturas recorrentes simples, tanto em conjuntos de dados artificiais, como visto em

Bengio, Simard e Frasconi (1994), como em dados reais, como os observados em Silva et al. (2023) e Li e Wu (2016). De modo que observa-se que a combinação inteligente de RNNs e CNNs em arquiteturas mais complexas continua a impulsionar avanços nas capacidades das redes neurais em lidar com dados do mundo real, em especial na classificação de sinais de ECG (STRODTHOFF et al., 2020).

4 Metodologia

O seguinte trabalho se dividiu em quatro etapas principais: a aquisição dos sinais a partir de uma base de dados, o processamento dos sinais, a elaboração de um modelo de rede neural e, por fim, a classificação dos sinais, na qual essa última inclui os ajustes na rede neural e diversas etapas de treinamento e validação a fim de otimizar o modelo. Um esboço do funcionamento geral do trabalho pode ser observado na Figura 10.

Figura 10 – Diagrama de funcionamento do trabalho.



Fonte: Autoral (2023).

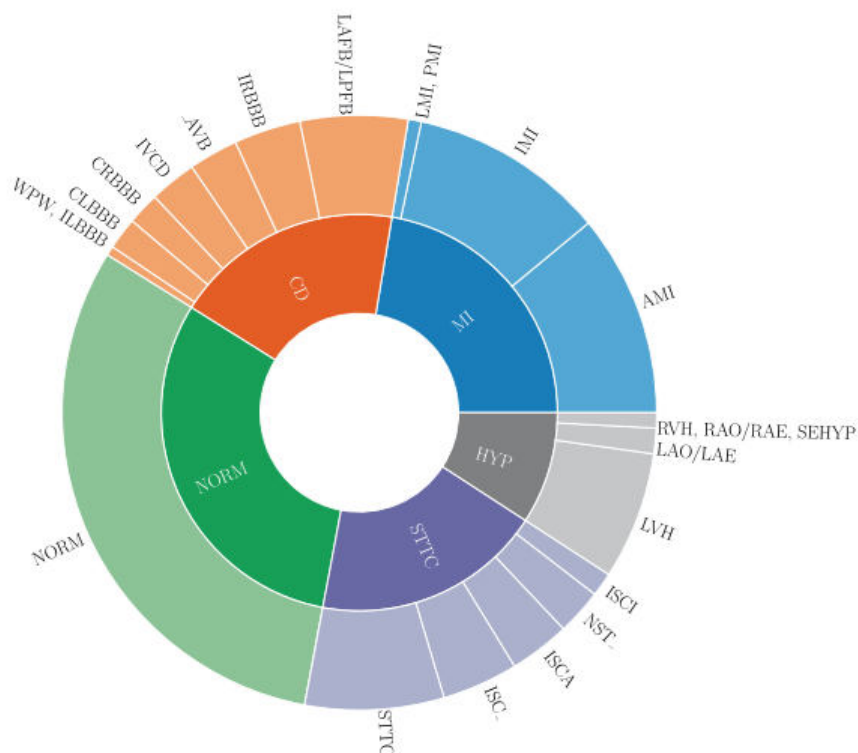
Nas seções seguintes será descrito como foi realizada a aquisição dos dados, o processamento dos sinais, apresentada a arquitetura da rede neural aplicada e a configuração utilizada nesse trabalho, apresentados os ajustes realizados na rede a fim de otimizar seus hiperparâmetros e, por fim, serão descritas as métricas utilizadas para avaliar o desempenho da rede.

4.1 Aquisição dos sinais

Nesta etapa, foi realizada a aquisição dos arquivos contendo as informações dos exames e sinais de ECG, oriundos da base de dados PTB-XL (WAGNER et al., 2020), uma base de dados pública apresentada em 2020 pelo Instituto Nacional de Metrologia da Alemanha (em alemão, *Physikalisch-Technische Bundesanstalt*, ou PTB), que também foi responsável pela curadoria de transferência de dados brutos. Essa base foi escolhida por ser a sucessora da base PTB e já ser amplamente reconhecida e difundida em pesquisas que envolvem desenvolvimento e testes de algoritmos para o processamento de sinais cardíacos, como nos trabalhos citados por Joloudari et al. (2022). Além disso, os sinais da base PTB-XL carregam diversos metadados com informações relevantes, como idade e sexo (STRODTHOFF et al., 2020).

O banco de dados compreende 21837 sinais de ECG de 12 derivações de 18885 pacientes (9820 homens e 9065 mulheres), com registros de 10 segundos de duração, e cobrindo idades de 0 a 95 anos, com mediana de 62 e intervalo interquartil de 22. Esses sinais, os quais possuem frequência de amostragem de 500 Hz, são de pacientes divididos entre 5 superclasses: Normal (NORM), Infarto do Miocárdio (MI), Distúrbio de Condução (CD), Hipertrofia (HYP) e Alteração do Segmento ST/T (STTC); e 24 subclasses, podendo o mesmo sinal estar em mais de uma subclasse a depender da patologia diagnosticada. A Figura 11 ilustra a divisão da base de dados em superclasses e subclasses.

Figura 11 – Divisão dos sinais do banco em suas superclasses e subclasses.



Fonte: (WAGNER et al., 2020).

O processo de obtenção dos ECG pelo PTB consistiu em duas etapas principais: a aquisição de dados e o pré-processamento. A aquisição de dados brutos foi realizada obtendo, para todos os sinais, o ECG para as 12 derivações por um intervalo de 10 segundos, com eletrodos de referência no braço direito e frequência de amostragem de 400 Hz. Os sinais foram avaliados por um cardiologista ou de forma automatizada (e posteriormente validadas por cardiologistas) e então cadastrados por um enfermeiro no banco de dados contendo os seus metadados associados, como sexo, idade, peso e altura, pseudonimizando nomes dos pacientes, enfermeiros e médicos e usando identificadores exclusivos a fim de preservação do anonimato. Posteriormente, os sinais são pré-processados, sendo convertidos para um formato binário com precisão de 16 bits a uma resolução de $1\mu\text{V}/\text{LSB}$ e, por fim, são sobreamostrados e subamostrados para frequência de 500 Hz e 100 Hz (WAGNER et al., 2020).

Os arquivos são disponibilizados na página da *PhysioNet* em formato DAT, contendo o identificador do sinal e os valores dele para cada derivação. Informações adicionais como gênero do paciente, idade e classe pertencente estão disponíveis em arquivos CSV (*Comma Separated Values*, em português: Valores separados por vírgula) à parte.

Foram lidos, então, os arquivos DAT contendo os sinais laudados por médicos para a derivação DI, particularmente os 7840 sinais pertencentes à superclasse NORM (Normal) e os 3485 sinais da superclasse MI (Infarto do Miocárdio), distribuindo-os em três conjuntos de dados estratificados: treino, validação e teste, seguindo uma proporção percentual de 70%, 15% e 15%, respectivamente.

4.2 Processamento dos sinais

Filtros de Resposta ao Impulso Finita (FIR, do inglês: *Finite Impulse Response*), também conhecidos como filtros *feedforward*, constituem uma classe de filtros cuja resposta ao impulso é limitada em extensão. A passagem pelo conjunto de coeficientes do filtro é unidirecional, o que implica que a saída do sistema é determinada exclusivamente pelo valor de entrada e seus atrasos correspondentes (OPPENHEIM; SCHAFER, 2013). Matematicamente, a saída $y[n]$ de um filtro FIR em resposta a uma entrada $x[n]$ pode ser expressa como:

$$y[n] = \sum_{k=0}^M h[k] \cdot x[n - k] \quad (4.1)$$

$$h[k] = \begin{cases} b_k, & k = 0, 1, \dots, M \\ 0, & \text{caso contrário} \end{cases} \quad (4.2)$$

em que b_k é o coeficiente do filtro $h[k]$ na posição k , e M é o número total de coeficientes no filtro.

Os filtros FIR possuem a capacidade de implementar diversas funções mediante ajustes em seus coeficientes, sendo categorizados como passa-baixa, passa-alta, passa-faixa, rejeita-faixa ou notch. Essas categorias representam diferentes respostas em frequência. A implementação de filtros FIR pode ocorrer de várias formas, dependendo das especificações do projeto. Neste trabalho, para remover artefatos e ruídos indesejados, como interferências da rede elétrica e ruídos musculares, foi realizada uma etapa de filtragem FIR passa faixas com frequências de corte [3, 45] Hz, como realizado por Costa (2017). Além disso, diferentes tamanhos do sinal de entrada foram testados, como feito por Feng et al. (2019), a fim de obter o tamanho ideal, como será melhor descrito na subseção 4.3.1.

4.3 Configuração da rede neural

Para o presente trabalho, foi implementada uma rede neural profunda, por meio da biblioteca Tensorflow (ABADI et al., 2016), combinando CNN e RNN, mais especificamente camadas de convoluções 1D e LSTM, trazendo uma abordagem híbrida para o processamento de dados eletrocardiográficos. O sinal de ECG, após o processamento, passa por 12 camadas de convolução e subamostragem, onde suas características espaciais são extraídas, e em seguida passa por uma LSTM, responsável por extrair suas características temporais. Por fim, tem-se uma camada MLP configurada para realizar uma tarefa de classificação binária. Um maior detalhamento da arquitetura da rede e suas camadas será feita a seguir.

A arquitetura da rede compreende um total de 15 camadas, sendo 9 camadas convolucionais, três camadas de subamostragem, uma camada de *dropout*, uma camada LSTM e, por último, uma camada MLP, semelhante ao modelo proposto por Feng et al. (2019). A Figura 12 ilustra a arquitetura da rede neural.

Nas camadas convolucionais unidimensionais foram utilizados números de saída da convolução iguais a 4, 8 e 16, respectivamente, com passo igual a 1 e tamanho do filtro (*kernel*) de convolução definido de acordo com os métodos descritos na subseção 4.3.1. Além disso, foi utilizada a função de ativação ReLU (Equação 4.3), função amplamente adotada para as camadas ocultas em função de suas derivadas serem, em geral, maiores em comparação com as de uma função sigmoide (Equação 4.4) e terem computação mais simples, além de aliviar o problema de *overfitting* e desvanecimento do gradiente e promover uma convergência mais rápida (GOODFELLOW; BENGIO; COURVILLE, 2016; RAMACHANDRAN; ZOPH; LE, 2017). A Figura 13 ilustra de forma gráfica as funções ReLU e sigmoide, respectivamente.

$$f(x) = \max(0, x) \quad (4.3)$$

Figura 12 – Arquitetura da rede neural elaborada.

```

Model: "model_conv1D"
-----
Layer (type)                Output Shape                Param #
-----
Conv4_1 (Conv1D)            (None, 2992, 4)            40
Conv4_2 (Conv1D)            (None, 2984, 4)            148
Conv4_3 (Conv1D)            (None, 2976, 4)            148
MaxPooling1D_1 (MaxPooling  (None, 595, 4)            0
1D)
Conv8_1 (Conv1D)            (None, 587, 8)             296
Conv8_2 (Conv1D)            (None, 579, 8)             584
Conv8_3 (Conv1D)            (None, 571, 8)             584
MaxPooling1D_2 (MaxPooling  (None, 114, 8)            0
1D)
Conv16_1 (Conv1D)           (None, 106, 16)            1168
Conv16_2 (Conv1D)           (None, 98, 16)             2320
Conv16_3 (Conv1D)           (None, 90, 16)             2320
MaxPooling1D_3 (MaxPooling  (None, 18, 16)            0
1D)
dropout (Dropout)           (None, 18, 16)            0
lstm (LSTM)                  (None, 16)                 2112
Dense (Dense)                (None, 1)                  17
=====

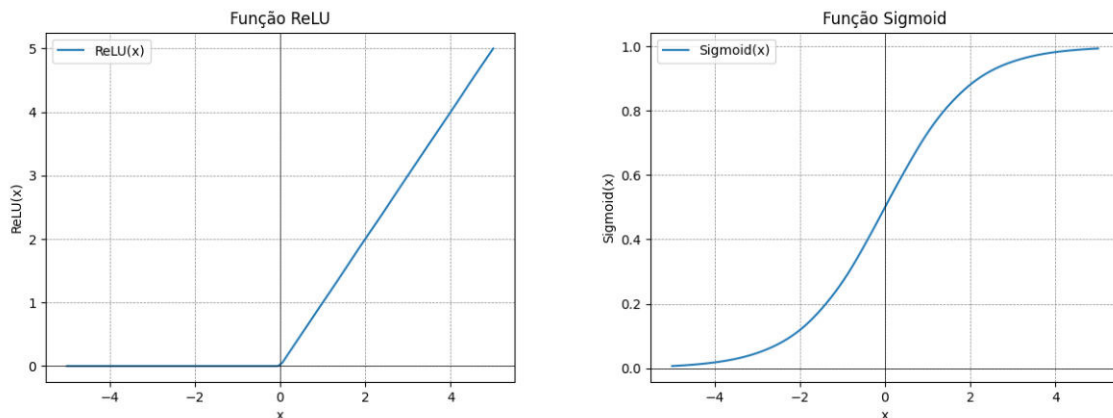
```

Fonte: Autoral (2023).

$$f(x) = \frac{1}{1 + e^{-x}} \quad (4.4)$$

As operações de subamostragem foram implementadas nas camadas 4, 8 e 12, utilizando um tamanho de *kernel* igual a 5. Essas camadas foram implementadas com o objetivo de reduzir a dimensionalidade dos parâmetros calculados, extraíndo as características relevantes para otimizar o processo de cálculo. Adicionalmente, foi aplicada uma camada de *dropout* entre a terceira camada de subamostragem e a camada LSTM, a fim de prevenir o *overfitting* através do desligamento de conjuntos aleatórios de neurônios, promovendo uma maior capacidade de generalização do modelo.

Após as camadas convolucionais e de *dropout* foi aplicada uma camada LSTM, a fim de obter as características temporais da saída da rede convolucional. Por fim, foi inserida uma camada MLP composta por um único neurônio, responsável por realizar a classificação binária dos sinais, entre saudável ou patológico, fazendo uso da função de ativação sigmoide (Equação 4.4).

Figura 13 – Função ReLU e função sigmoide.

Fonte: Autoral (2023).

4.3.1 Seleção de hiperparâmetros da rede neural

Redes neurais contêm hiperparâmetros, que são configurações externas ao modelo e que determinarão a operação da rede. Dentre os hiperparâmetros estão definidos o número de camadas ocultas, o número de neurônios de cada camada oculta, a função de ativação utilizada em cada camada, a taxa de aprendizagem, função de perda, taxa de *dropout*, entre outros.

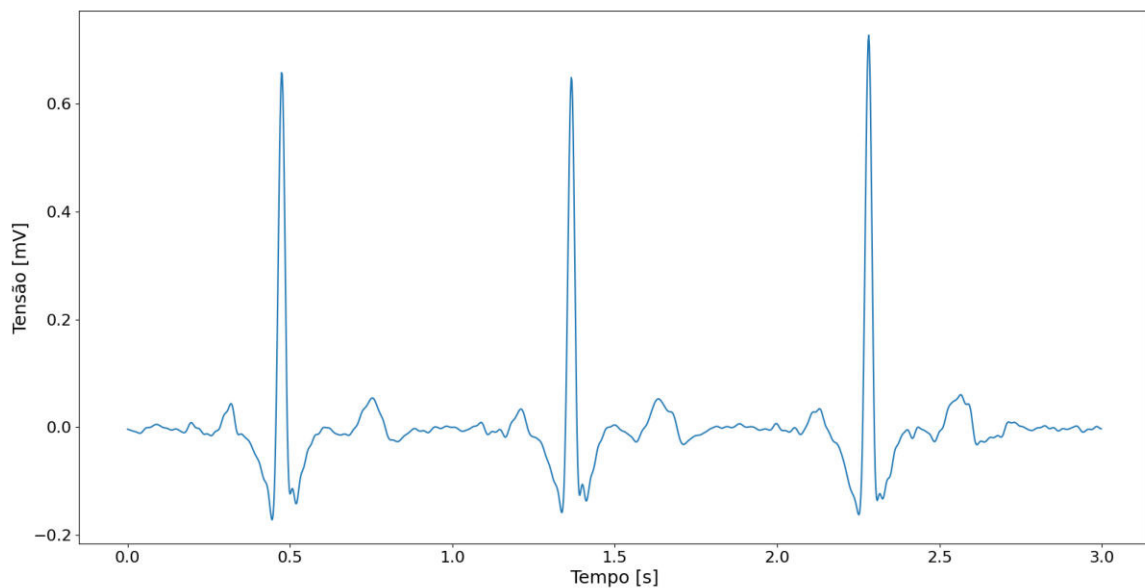
A escolha dos hiperparâmetros afeta diretamente o desempenho do treinamento e teste da rede, de forma que um número alto de camadas ou neurônios em uma rede pode aumentar sua aptidão para lidar com os dados do conjunto de treinamento, porém, pode comprometer sua capacidade de generalização. Além disso, em redes neurais profundas, a acumulação das derivadas de uma função de ativação com valores baixos pode resultar no desvanecimento do gradiente, o que interrompe o processo de treinamento (ROODS-CHILD; SARDIÑAS; WILL, 2020), impedindo a atualização dos parâmetros da rede, algo que não ocorre quando se utiliza uma função de ativação com uma derivada de valor constante.

Dessa forma, é necessário testar combinações de diferentes valores desses hiperparâmetros para colaborar com a geração de modelos capazes de terem capacidade de generalização satisfatória para dados não apresentados no treinamento. Esses testes foram realizados por meio do método *Bayesian Optimization Tuner* do *KerasTuner* (CHOLLET et al., 2015).

Uma preparação do conjunto de dados para o *Bayesian Optimization Tuner* foi realizada, a qual utilizou 10% dos sinais dos conjuntos de treino e validação. O *tuner* operou por meio de 200 iterações, testando diversas combinações de taxas de *dropout*, tamanho do *kernel* de convolução, taxas de aprendizado e otimizadores, incluindo o Adam, SGD e RMSprop. Após a identificação dos melhores hiperparâmetros, a rede neural passou por

cinco rodadas de treinamento e validação usando a configuração ótima de hiperparâmetros, a fim de se obter suas médias e desvios padrão. Por fim, cada teste foi repetido, variando o tamanho dos sinais de entrada da rede entre 3, 6 e 9 segundos, equivalentes a, respectivamente, 3, 6 e 9 ciclos cardíacos, aproximadamente, como observado na Figura 14. Com isso, buscou-se normalizar o tamanho das entradas da rede, bem como determinar o comprimento mais adequado para a otimização do seu desempenho.

Figura 14 – Amostra de sinal de ECG de 3 s.



Fonte: Autoral (2023).

Além disso, foram utilizadas as ferramentas *EarlyStopping* e *ReduceLROnPlateau* do *Keras* (CHOLLET et al., 2015), que têm por objetivo aprimorar o treinamento da rede neural. O *EarlyStopping* monitora o desempenho da rede durante o treinamento e interrompe o processo caso não haja melhorias significativas, evitando o *overfitting*. Já o *ReduceLROnPlateau* ajusta a taxa de aprendizado após um determinado número de épocas sem melhorias no aprendizado da rede, ou seja, quando o modelo atinge um platô, o que permite uma convergência mais eficiente. Essas estratégias contribuem para a estabilidade e eficiência do processo de treinamento da rede neural.

4.4 Métricas de desempenho

Por se tratar de uma aplicação que utilizou um conjunto de dados muito desbalanceado, em que apenas 30% dos sinais são patológicos, a Acurácia (Acc) não será utilizada como a principal métrica de desempenho, pois a mesma representa uma avaliação incompleta acerca do desempenho da rede (CASTRO; BRAGA, 2011). Portanto, para avaliar o seu desempenho, também foram utilizadas as métricas Precisão (Prec), Sensibilidade (Sen) e F1-Score.

A Acurácia indica a performance geral do modelo, de forma que quantifica a porcentagem de quantas classificações foram realizadas corretamente. Sua definição matemática pode ser observada na Equação 4.5, na qual o TP diz respeito aos verdadeiros positivos, TN aos verdadeiros negativos, FP aos falsos positivos e FN aos falsos negativos.

$$\text{Acc} = \frac{TP + TN}{TP + FP + TN + FN} \quad (4.5)$$

A Precisão consiste na porcentagem de acerto dos sinais classificados como positivo, como observado na Equação 4.6.

$$\text{Prec} = \frac{TP}{TP + FP} \quad (4.6)$$

A Sensibilidade por sua vez consiste em uma métrica que indica a tendência da rede em classificar sinais como positivos e pode ser entendida a partir da Equação 4.7.

$$\text{Sen} = \frac{TP}{TP + FN} \quad (4.7)$$

O F1-Score pode ser entendido como a média harmônica entre precisão e sensibilidade e sua definição matemática pode ser observada na Equação 4.8.

$$\text{F1} = 2 \cdot \frac{\text{Prec} \cdot \text{Sen}}{\text{Prec} + \text{Sen}} \quad (4.8)$$

A Loss, ou função de perda, é uma métrica que quantifica o quão distantes as predições do modelo estão dos valores reais. Em problemas de classificação binária, como neste trabalho, a *Cross-Entropy* é frequentemente utilizada como função de perda. Seu objetivo é minimizar a divergência entre as probabilidades preditas e as probabilidades reais dos exemplos de treinamento. A expressão matemática da *Cross-Entropy Loss* é dada pela Equação 4.9:

$$\text{Loss} = -\frac{1}{N} \sum_{i=1}^N [y_i \cdot \log(\hat{y}_i) + (1 - y_i) \cdot \log(1 - \hat{y}_i)] \quad (4.9)$$

na qual N é o número total de exemplos, y_i é a verdade (1 se o exemplo é positivo, 0 se é negativo), e \hat{y}_i é a predição do modelo para o exemplo i .

Por fim, a AUC (*Area Under the Curve*) é uma métrica que avalia a capacidade do modelo de distinguir entre classes. No contexto de classificação binária, a AUC representa a área sob a curva ROC (*Receiver Operating Characteristic*), que plota a Sensibilidade contra a Taxa de Falsos Positivos. A AUC varia de 0 a 1, onde 0,5 indica um desempenho similar ao acaso, e 1 indica um desempenho perfeito, portanto, quanto maior a AUC, melhor o desempenho do modelo em discriminar entre classes.

5 Resultados

Nessa seção serão apresentados os resultados e comentários acerca do desempenho da rede classificadora de sinais de ECG com base nos tópicos apresentados no Capítulo 4. Para isso, serão apresentados os resultados das métricas de validação para identificação do infarto do miocárdio, além do conjunto de hiperparâmetros ótimo para cada comprimento de sinal. Deve-se destacar que o F1-Score é a principal métrica a ser observada, por ser a média harmônica entre sensibilidade e precisão, entretanto, em se tratando de um sistema de identificação de patologias, a sensibilidade deve receber atenção especial, sendo preferível uma alta sensibilidade.

A Tabela 1 informa os hiperparâmetros ótimos obtidos para cada comprimento de sinal de entrada, enquanto que as Tabelas 2 e 3 apresentam as métricas de classificação do conjunto de dados de validação e o seu desvio padrão, respectivamente, devendo-se destacar que essas métricas foram obtidas por meio de 5 rodadas de treinamento e validação após a obtenção dos hiperparâmetros ótimos, onde as médias e desvio-padrão foram calculadas para as 5 rodadas de validação.

Tabela 1 – Hiperparâmetros ótimos para os diferentes comprimentos de sinais de ECG. (*Kernel* diz respeito ao tamanho do filtro de convolução, *LSTM* corresponde ao tamanho da saída da camada correspondente e *LR* diz respeito à taxa de aprendizado.)

	Dropout	Kernel	LSTM	Otimizador	LR
3 s	0,45	7	8	Adam	0.001
6 s	0,50	9	8	RMSprop	0.001
9 s	0,45	9	12	Adam	0.001

Tabela 2 – Resultados das métricas para os diferentes comprimentos de sinais de ECG.

	Loss	Acc(%)	Prec(%)	Sen(%)	F1(%)	AUC
3 s	0,3711	84,85	76,45	74,41	75,36	0,8920
6 s	0,3598	85,32	80,44	70,25	74,88	0,8926
9 s	0,3656	84,94	77,75	72,69	75,03	0,8925

Observando a Tabela 2, nota-se que o comprimento de sinal que apresentou melhor acurácia foi o de 6 segundos, com 85,32%, que também apresentou o melhor valor de *Loss* e precisão, 0,3598 e 80,44%, respectivamente. Entretanto, nele foram observados os piores valores de F1-Score e sensibilidade, 74,88% e 70,25% respectivamente, o que não é desejado num sistema de identificação de patologias, sendo preferível que a rede tenha

maior tendência para categorizar sinais como falsos positivos do que falsos negativos, pois dessa forma o paciente permanece sob cuidados médicos.

O comprimento de sinal de 3 segundos, por outro lado, apresentou as maiores métricas de F1-Score e sensibilidade, 75,36% e 74,71% respectivamente, sendo mais adequado para a aplicação desejada por conta de sua sensibilidade 4,14% mais elevada que a do comprimento de 6 segundos. Além disso, seus valores de acurácia e AUC são semelhantes aos de 6 segundos, de forma que sua desvantagem em relação aos sinais de 6 segundos se encontram na *Loss* mais elevada e na precisão 3,99% menor.

Por fim, o comprimento de sinal de 9 segundos apresentou métricas com valores intermediários em relação aos comprimentos de sinal anteriores, com 84,94% de acurácia, 77,75% de precisão, 72,69% de sensibilidade e F1-Score de 75,03%. Além disso, seus valores de *Loss* e AUC se mantiveram semelhantes aos sinais anteriores.

Tabela 3 – Resultados dos desvios padrão para os diferentes comprimentos de sinais de ECG.

	Loss	Acc(%)	Prec(%)	Sen(%)	F1(%)	AUC
3 s	0,0152	0,79	2,81	2,71	1,08	0,0063
6 s	0,0138	0,67	3,79	3,15	0,44	0,0030
9 s	0,0134	0,69	3,19	3,33	0,88	0,0051

Observando a Tabela 3 pode-se notar que as métricas com maiores desvios são a precisão e sensibilidade, especialmente nos comprimentos de sinal de 6 e 9 segundos. Além disso, o comprimento de 3 segundos se destacou com um desvio no F1-Score consideravelmente superior ao dos demais comprimentos de sinal, apresentando também o maior valor de desvio padrão para *Loss*, acurácia e AUC.

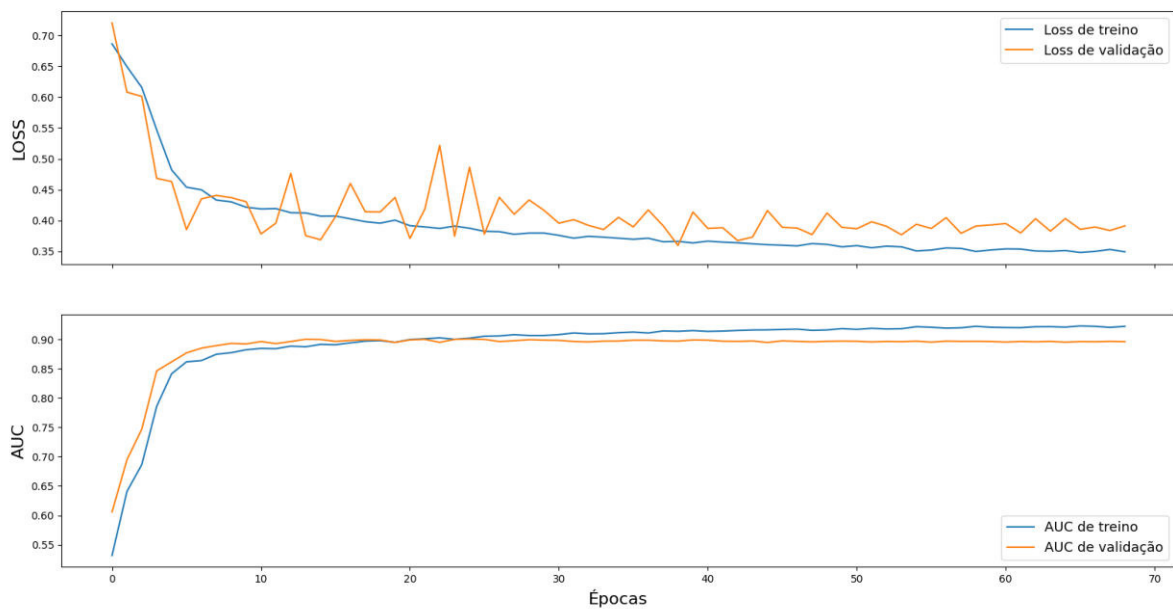
5.1 Teste da rede

Por conta do F1-Score e sensibilidade mais elevados, para esta etapa, foi selecionado o comprimento de sinal de ECG de 3 segundos para a classificação do conjunto de dados de teste. Para isso, foram utilizados os hiperparâmetros ótimos para esse comprimento de sinal, tal como visto na Tabela 1. Além disso, a proporção entre sinais saudáveis e patológicos foi mantida igual à dos conjuntos de dados de treinamento e validação.

A evolução das métricas *Loss* e AUC durante o treino final da rede podem ser observadas na Figura 15, onde a curva da função de perda (*Loss*) no conjunto de treino oferece uma indicação do quão efetivamente o modelo está aprendendo durante o treinamento, enquanto a curva da função de perda no conjunto de validação indica a capacidade de generalização do modelo, essencial para lidar com dados não vistos, como os do conjunto de teste. Observa-se que a curva de *Loss* do conjunto de treino exibe um comportamento

esperado, diminuindo seu valor até atingir um ponto de estabilização. Já a *Loss* do conjunto de validação, apesar de apresentar variações significativas no início do treinamento, demonstra uma estabilização após aproximadamente 50 épocas, mantendo um intervalo de generalização, ou seja, distância entre as curvas, constante após isso. Em relação à métrica AUC, observa-se que suas curvas apresentam um comportamento dentro do esperado, indicando uma discriminação eficaz entre as classes, e mantendo um intervalo constante também após aproximadamente 50 épocas, sugerindo que o modelo atingiu uma estabilidade na capacidade de distinção entre as classes, mantendo uma distância constante entre os conjuntos de treino e validação ao longo do treinamento.

Figura 15 – Evolução da função de perda (*Loss*) e AUC durante o treino do modelo final da rede.



Fonte: Autoral (2023).

O desempenho do teste da rede pode ser avaliado através das métricas observadas na Tabela 4 e pela matriz de confusão observada na Figura 16. Observa-se na Tabela 4 que todas as métricas foram decrescidas, sendo a precisão a métrica que apresentou a maior queda, 1,98%, ao lado do F1-Score, que apresentou queda de 1,03%. As demais métricas não apresentaram quedas superiores a 1%. Além disso, comparando com a Tabela 3, observa-se que as métricas de teste apresentaram valores dentro do desvio previsto.

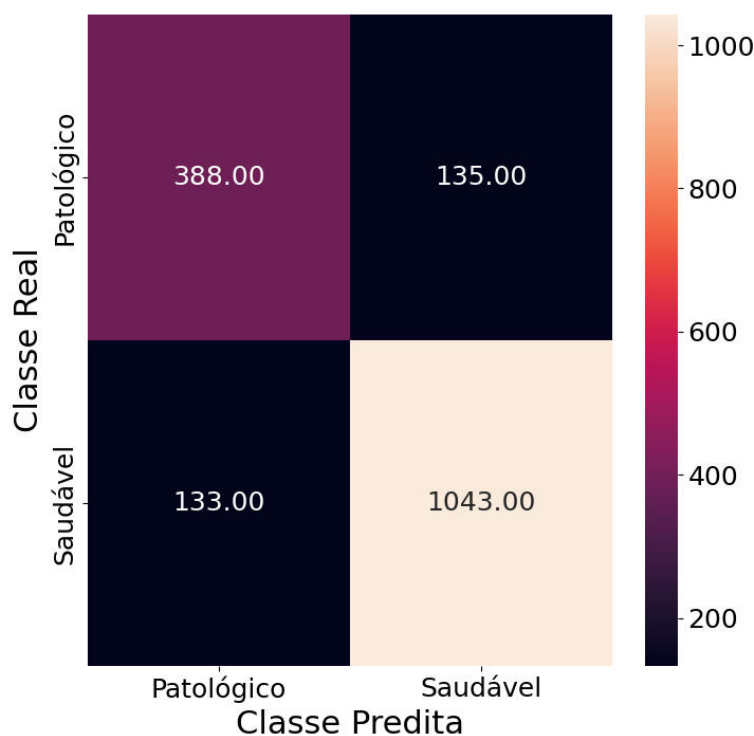
Tabela 4 – Resultado do teste da rede.

	Acc(%)	Prec(%)	Sen(%)	F1(%)
3 s	84,23	74,47	74,19	74,33

Pode-se observar, também, a matriz de confusão da classificação dos sinais de ECG na Figura 16, onde é possível observar a acurácia de 84,23% na forma da diagonal

principal da matriz, além de ser possível notar a similaridade entre os valores de precisão e sensibilidade, de forma que os erros da rede estão bem distribuídos entre a diagonal secundária.

Figura 16 – Matriz de confusão para classificação de ECG.



Fonte: Autoral (2023).

5.2 Discussões

Foi observado, portanto, que fazendo uso de apenas uma derivação eletrocardiográfica, especificamente a derivação DI, o comprimento do sinal teve pouco impacto no F1-Score, acurácia e AUC, sendo a precisão e sensibilidade as métricas que apresentaram maior variação. Além disso, os melhores valores de F1-Score e sensibilidade se fizeram presentes no comprimento de sinal de 3 segundos, o que o levou a ser o escolhido para a aplicação do conjunto de dados de teste.

O trabalho apresentou métricas de acurácia 9,96% inferiores às observadas em He et al. (2021), entretanto deve-se ressaltar a limitação no número de derivações utilizadas, onde no presente trabalho foi utilizada apenas a DI enquanto que no modelo proposto por He et al. (2021) todas as 12 derivações foram aplicadas. Além disso, comparando com os resultados vistos em Vieira et al. (2022), o modelo de rede neural híbrido foi capaz de alcançar métricas consideravelmente superiores em relação ao modelo MLP, onde se

alcançou acurácia 13,92% e sensibilidade 4,43% superiores, o que sugere um desempenho melhor da arquitetura proposta em relação à abordagem com MLP.

6 Conclusão

A incidência global de doenças cardiovasculares impacta profundamente a qualidade de vida de inúmeras pessoas, destacando a urgência de soluções eficientes para acelerar diagnósticos. Com o aumento de estudos em processamento e classificação de sinais cardíacos, impulsionados por repositórios públicos de dados, o presente trabalho busca contribuir com a redução do número de fatalidades, diminuição de tempos de espera e aprimoramento da triagem de casos suspeitos. Esses avanços têm o potencial de melhorar significativamente a saúde cardiovascular e a qualidade de vida dos afetados por tais condições.

O presente trabalho teve por objetivo a implementação e teste de uma arquitetura de rede neural híbrida, combinando CNN e RNN, para a classificação de sinais de ECG de derivação única, a fim de descobrir o comprimento de sinal ideal para entrada da rede, bem como o conjunto de hiperparâmetros que permite a rede alcançar o seu melhor resultado, permitindo assim criar um sistema possível de ser embarcado em dispositivos vestíveis, promovendo um rápido diagnóstico para casos de condições de risco. Para isso, três diferentes comprimentos de sinal foram comparados a partir das suas métricas de classificação.

Os resultados obtidos na etapa de validação sugeriram que comprimentos de sinal menores apresentam maior sensibilidade e F1-Score, métricas preferíveis em se tratando de um sistema de identificação de patologias, por assegurar o tratamento médico do paciente. Além disso, o aumento do comprimento do sinal não promoveu melhorias significativas para o desempenho da rede, não sendo possível testar com comprimentos superiores a 9 segundos por limitação do banco de dados utilizado.


Por fim, sugestões para trabalhos futuros incluem a adaptação da arquitetura da rede para trabalhar com comprimentos de sinal inferiores a 3 segundos, a fim de verificar se há melhoria nas métricas de classificação. Além disso, sugere-se a adaptação da arquitetura da rede a fim de incorporar as informações de idade e sexo dos pacientes presentes nos metadados, visando aprimorar a capacidade de classificação da rede e assim alcançar resultados mais robustos.

REFERÊNCIAS

- ABADI, M. et al. Tensorflow: A system for large-scale machine learning. In: *12th {USENIX} symposium on operating systems design and implementation ({OSDI} 16)*. [S.l.: s.n.], 2016. p. 265–283.
- BENGIO, Y.; SIMARD, P.; FRASCONI, P. Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*, IEEE, 1994.
- BLAUS, B. Medical gallery of blausen medical 2014. *Wiki J. Med*, v. 1, n. 10, 2014.
- BRAUNWALD, E. et al. *Braunwald's Heart Disease: A Textbook of Cardiovascular Medicine. 11th edition*. [S.l.]: Elsevier Science, 2019.
- CARDIOLOGIA, S. B. de. *Diretriz de interpretação de eletrocardiograma de repouso*, [S.l.], 2003.
- CASTRO, C. L. d.; BRAGA, A. P. Aprendizado supervisionado com conjuntos de dados desbalanceados. *SBA: Controle Automação Sociedade Brasileira de Automatica*, Sociedade Brasileira de Automática, v. 22, n. 5, p. 441–466, Sep 2011. ISSN 0103-1759. Disponível em: <<https://doi.org/10.1590/S0103-17592011000500002>>.
- CHOLLET, F. et al. *Keras*. 2015. <<https://keras.io>>.
- COSTA, C. *Identificação do Infarto do Miocárdio por meio da Análise do Cardiograma Vetorial e da Reconstrução do Espaço de Fases*. Dissertação (Mestrado) — Instituto Federal da Paraíba, 2017.
- FELDMAN, J.; GOLDWASSER, G. Eletrocardiograma: recomendações para a sua interpretação. *Revista da Sociedade de Cardiologia do Estado do Rio de Janeiro*, Sociedade Brasileira de Cardiologia, v. 17, n. 4, p. 251–256, 2004.
- FENG, K. et al. Myocardial infarction classification based on convolutional neural network and recurrent neural network. *Applied Sciences*, v. 9, n. 9, 2019. ISSN 2076-3417. Disponível em: <<https://www.mdpi.com/2076-3417/9/9/1879>>.
- GAZIANO, T. et al. Cardiovascular disease. In: *Disease Control Priorities in Developing Countries. 2nd edition*. [S.l.]: The International Bank for Reconstruction and Development/The World Bank, 2006.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep Learning*. [S.l.]: MIT Press, 2016. <<http://www.deeplearningbook.org>>.
- HAGIWARA, Y. et al. Computer-aided diagnosis of atrial fibrillation based on ecg signals: A review. *Information Sciences: Informatics and Computer Science Intelligent Systems Applications*, Elsevier Inc., v. 467, p. 99–114, 2018.
- HAMPTON, J. *150 ECG Casos Clínicos*. [S.l.]: Elsevier Brasil, 2014.
- HAYKIN, S. O. *Neural Networks and Learning Machines/Simon Haykin*. 3rd. ed. [S.l.]: Pearson Education, Inc., 2009.

- HE, Z. et al. Mfb-lann: A lightweight and updatable myocardial infarction diagnosis system based on convolutional neural networks and active learning. *Computer methods and programs in biomedicine*, Elsevier B.V, Ireland, v. 210, p. 106379–106379, 2021. ISSN 0169-2607.
- HOCHREITER, S.; BENGIO, Y.; FRASCONI, P. Gradient flow in recurrent nets: the difficulty of learning long-term dependencies. *IEEE Press*, IEEE, 2001.
- HSU, H. *Teoria e problemas de sinais e sistemas*. [S.l.]: Bookman, 2004.
- JOLOUDARI, J. H. et al. Application of artificial intelligence techniques for automated detection of myocardial infarction: a review. *Physiological Measurement*, IOP Publishing, v. 43, n. 8, p. 08TR01, aug 2022. Disponível em: <<https://dx.doi.org/10.1088/1361-6579/ac7fd9>>.
- KARAKAYA, M. *1 Dimensional Convolution (Conv1D) for Regression: Predict house prices*. 2021. Disponível em: <<https://kmkarakaya.medium.com/>>.
- LI, X.; WU, X. Long short-term memory based convolutional recurrent neural networks for large vocabulary speech recognition. *arXiv.org*, Cornell University Library, arXiv.org, Ithaca, 2016. ISSN 2331-8422.
- MOZAFFARIAN, D. et al. Heart disease and stroke statistics—2016 update. *Circulation*, v. 133, n. 4, p. e38–e360, 2016. Disponível em: <<https://www.ahajournals.org/doi/abs/10.1161/CIR.0000000000000350>>.
- NISA, E. C.; KUAN, Y.-D. Comparative assessment to predict and forecast water-cooled chiller power consumption using machine learning and deep learning algorithms. *Sustainability*, v. 13, n. 2, 2021. ISSN 2071-1050. Disponível em: <<https://www.mdpi.com/2071-1050/13/2/744>>.
- OPPENHEIM, A.; SCHAFER, R. *Processamento em tempo discreto de sinais*. 3rd. ed. [S.l.]: São Paulo: Pearson, 2013.
- ORGANIZATION, W. H. *Cardiovascular diseases (CVDs)s*. 2017. Disponível em: [https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)). Acesso em: Fevereiro de 2021.
- PASTORE, C. et al. Iii diretrizes da sociedade brasileira de cardiologia sobre análise e emissão de laudos eletrocardiográficos. *Revista da Sociedade Brasileira de Cardiologia*, Sociedade Brasileira de Cardiologia, v. 106, n. 4, 2016.
- PETMEZAS, G. et al. State-of-the-art deep learning methods on electrocardiogram data: Systematic review. *JMIR Med Inform*, aug 2022.
- PIEGAS, L. et al. V diretriz da sociedade brasileira de cardiologia sobre tratamento do infarto agudo do miocárdio com supradesnível do segmento st. *Arquivos Brasileiros de Cardiologia*, Arquivos Brasileiros de Cardiologia, Avenida Marechal Câmara 160-330, Centro, Rio de Janeiro, RJ, v. 105, n. 2, p. 1–105, 2015.
- RAMACHANDRAN, P.; ZOPH, B.; LE, Q. V. Searching for activation functions. *arXiv preprint arXiv:1710.05941*, 2017.
- REIS, H. et al. *ECG: manual prático de eletrocardiograma*. [S.l.]: Atheneu, 2013.

- ROODSCHILD, M.; SARDIÑAS, J. G.; WILL, A. A new approach for the vanishing gradient problem on sigmoid activation. *Progress in Artificial Intelligence*, Springer, v. 9, n. 4, p. 351–360, 2020.
- SCHWARZ, L. Artigo de revisão: Eletrocardiograma. *Rev. Ilha Digital*, 2009.
- SILVA, W. et al. Classificação de sinais de eeg para diagnóstico do transtorno do espectro autista utilizando cnn e rnn. *XLI Simpósio Brasileiro de Telecomunicações e Processamento de Sinais*, 2023.
- STRODTHOFF, N. et al. Deep learning for ecg analysis: Benchmarks and insights from ptb-xl. *arXiv preprint arXiv:2004.13701*, 2020.
- THYGESEN, K. et al. Universal definition of myocardial infarction. *circulation*, Am Heart Assoc, v. 116, n. 22, p. 2634–2653, 2007.
- VIEIRA, V. J. D. et al. Classificação do infarto de miocárdio a partir da extração de parâmetros de espaços e fase com o uso dos descritores de haralick. *XL Simpósio Brasileiro de Telecomunicações e Processamento de Sinais*, 2022.
- WAGNER, P. et al. Pt看xl, a large publicly available electrocardiography dataset. *Scientific Data*, Nature Publishing Group, v. 7, n. 1, p. 1–15, 2020.
- WHITE, H. D.; CHEW, D. P. Acute myocardial infarction. *The Lancet*, Elsevier, v. 372, n. 9638, p. 570–584, 2008.

	INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DA PARAÍBA
	Campus João Pessoa
	Av. Primeiro de Maio, 720, Jaguaribe, CEP 58015-435, Joao Pessoa (PB)
	CNPJ: 10.783.898/0002-56 - Telefone: (83) 3612.1200

Documento Digitalizado Ostensivo (Público)

Trabalho de Conclusão de Curso

Assunto:	Trabalho de Conclusão de Curso
Assinado por:	Gabriel Gomes
Tipo do Documento:	Anexo
Situação:	Finalizado
Nível de Acesso:	Ostensivo (Público)
Tipo do Conferência:	Cópia Simples

Documento assinado eletronicamente por:

- **Gabriel Diniz Gomes, ALUNO (20191610003) DE BACHARELADO EM ENGENHARIA ELÉTRICA - JOÃO PESSOA**, em 27/12/2023 12:33:03.

Este documento foi armazenado no SUAP em 27/12/2023. Para comprovar sua integridade, faça a leitura do QRCode ao lado ou acesse <https://suap.ifpb.edu.br/verificar-documento-externo/> e forneça os dados abaixo:

Código Verificador: 1039388

Código de Autenticação: b66d19df06

