

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

DAVI CAUASSA LEÃO

DETECÇÃO E CLASSIFICAÇÃO AUTOMÁTICA DE ARRITMIAS CARDÍACAS  
EM SINAIS DE ELETROCARDIOGRAMA UTILIZANDO REDES  
CONVOLUTIVAS 2D

MANAUS

2024

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

DAVI CAUASSA LEÃO

DETECÇÃO E CLASSIFICAÇÃO AUTOMÁTICA DE ARRITMIAS  
CARDÍACAS EM SINAIS DE ELETROCARDIOGRAMA UTILIZANDO  
REDES CONVOLUTIVAS 2D

Dissertação apresentada ao Curso de Mestrado em Engenharia Elétrica, área de concentração Controle e Automação de Sistemas e linha de pesquisa Reconhecimento de Padrões e Otimização do Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas.

Orientadora: Profa. Dra. Marly Guimarães Fernandes Costa

Coorientador: Prof. Dr. Cícero Ferreira Fernandes Costa Filho

MANAUS

2024

## Ficha Catalográfica

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

L437d Leão, Davi Cauassa  
Detecção e classificação automática de arritmias cardíacas em sinais de eletrocardiograma utilizando redes convolutivas 2d / Davi Cauassa Leão. 2024  
96 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. ECG- eletrocardiograma. 2. Arritmia. 3. Aumento de dados. 4. rede neural Convolutional. I. Costa, Marly Guimarães Fernandes. II. Universidade Federal do Amazonas III. Título



Poder Executivo  
Ministério da Educação  
Universidade Federal do Amazonas  
Faculdade de Tecnologia  
Programa de Pós-graduação em Engenharia Elétrica

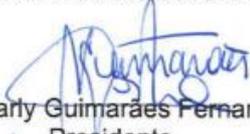
**DAVI CAUASSA LEÃO**

**DETECÇÃO E CLASSIFICAÇÃO AUTOMÁTICA DE ARRITMIAS  
CARDÍACAS EM SINAIS DE ELETROCARDIOGRAMA UTILIZANDO  
REDES CONVOLUTIVAS 2D.**

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 09 de fevereiro de 2024.

**BANCA EXAMINADORA**

  
Prof.<sup>a</sup>. Dra. Marly Guimarães Fernandes Costa  
Presidente  
Universidade Federal do Amazonas

  
Prof. Dr. Francisco de Assis Pereira Januário, Membro  
Universidade Federal do Amazonas

  
Prof. Dr. Frederico da Silva Pinagé, Membro  
Universidade Federal do Amazonas



Pós-Graduação em Engenharia Elétrica.  
Av. General Rodrigo Octávio Jordão Ramos, nº 3.000 - Campus  
Universitário, Setor Norte - Coroado, Pavilhão do CETELI.  
Fone/Fax (92) 99271-8954 Ramal:2607. E-mail: ppgee@ufam.edu.br

## **AGRADECIMENTOS**

Agradeço a Deus pela minha vida e por me permitir superar todos os obstáculos encontrados ao longo da realização deste trabalho.

A minha linda esposa, Acsa Leão, por estar sempre ao meu lado, apoiando-me e abdicando do seu tempo comigo quando precisei estudar, também por toda a estrutura que pode me fornecer, pelos cuidados, por seu amor e pelas palavras de superação que sempre me ajudaram. Amo muito você e que nossos dias sejam eternos.

Ao meu filho, Marcelo Ruan, pela paciência demonstrada em várias ocasiões em que não tivemos tempo de qualidade devido aos estudos.

Aos meus pais e irmãos, por estarem ao meu lado e me apoiarem nos momentos mais difíceis.

A Safira Leão, por 15 anos ao meu lado, sempre me acompanhando nos estudos.

Aos meus orientadores, Dra. Marly Guimarães Fernandes Costa e Prof. Dr. Cícero Ferreira Fernandes Costa Filho, pela orientação, paciência e pela confiança que diversas vezes depositaram em mim e por todo conhecimento e experiência que me passaram.

Aos meus amigos, João Victor, Mikaela Serrão, Neilson e Matheus de Paula pela ajuda durante todo o processo de pesquisa e desenvolvimento.

À instituição CAPES pelo fornecimento da bolsa que possibilitou minha dedicação exclusiva ao desenvolvimento da dissertação. À Universidade Federal do Amazonas e, em particular, ao Centro de Tecnologia Eletrônica e da Informação – CETELI – pela concessão de toda a infraestrutura para a realização deste trabalho.

Esta dissertação, realizada no âmbito do Projeto Samsung - UFAM de Ensino e Pesquisa (SUPER), de acordo com o Artigo 39 do Decreto nº 10.521/2020, foi financiada pela Samsung Eletrônica da Amazônia Ltda, nos termos da Lei Federal nº 8.387/1991, através do convênio 001/2020 firmado com a UFAM e FAEPI, Brasil.

## RESUMO

O ECG é o método diagnóstico mais utilizado para identificar e classificar arritmias cardíacas, infarto do miocárdio e doença arterial coronariana. A maioria dos artigos apresentados na literatura para classificação de arritmias em eletrocardiogramas (ECG) utiliza redes convolucionais 1D. Essas redes exploram apenas as relações entre vizinhos à esquerda e à direita. Redes convolucionais 2D, por outro lado, exploram adicionalmente as relações entre vizinhos acima e abaixo, permitindo o estabelecimento de relações mais complexas entre os dados. Este estudo propõe um método para a classificação de arritmias usando redes convolucionais 2D. As amostras do sinal de ECG (1D) é convertido em uma imagem de intensidade (2D). Para o treinamento da CNN, uma nova técnica de aumento de dados é proposta utilizando janelas deslizantes. Valores de aumento de dados de até 11 vezes são alcançados. Três arquiteturas de CNN e dois métodos de otimização (SGDM e ADAM) são avaliados. O melhor resultado para a classificação de 13 arritmias, com uma precisão global de 95,80%, é obtido com o otimizador SGDM e aumento de dados.

**Palavras-chave:** ECG, arritmia, aumento de dados, rede neural Convolucional.

## ABSTRACT

The ECG is the most widely used diagnostic method for identifying and classifying cardiac arrhythmias, myocardial infarction and coronary artery disease. Most articles presented in the literature for the classification of arrhythmias in ECG use 1D convolutional networks. These networks only explore relationships between left and right neighbors. On the other hand, 2D convolutional networks (CNN) also explore relationships between upstream and downstream neighbors, allowing more complex relationships to be established among the data. This study proposes a method for arrhythmia classification using 2D convolutional networks using CNN. The 1D ECG signal is converted into an intensity image. For CNN training, a new data augmentation technique is proposed using sliding windows. Data augmentation values up to 11 times are achieved. Three CNN architectures and two optimization methods (SGDM, and ADAM) are evaluated. The best result for the classification of 13 arrhythmias, with an overall accuracy of 95.80%, is obtained with the SGDM optimizer and data augmentation.

**Keywords:** *ECG, arrhythmia, data augmentation, convolutional neural network.*

## LISTAS DE FIGURAS

Figura 1. Extração de imagens direto do ECG .....	32
Figura 2. Transformação de sinais unidimensionais em espectrograma através da transformada <i>wavelet</i> contínua. ....	33
Figura 3. Anatomia do coração .....	36
Figura 4. Fluxos sanguíneos dentro do coração. ....	36
Figura 5. Caracterização esquemática do ECG com todas as suas ondas (P, Q, R, S, T e U) e seus intervalos (PR, QT e ST). ....	37
Figura 6. Eletrocardiograma. ....	38
Figura 7. Representação do papel milimetrado do ECG com a duração em segundos na linha horizontal e a amplitude em milivolts na linha vertical .....	38
Figura 8. Leitura de sinais elétricos no método <i>on the person</i> .....	39
Figura 9. Pontos de coleta de frequências cardíacas em um ECG.....	40
Figura 10. Trecho de um exame de ECG contendo a arritmia batida de bloqueio de ramo direito (R).....	41
Figura 11. Trecho de um exame de ECG contendo a arritmia contração ventricular prematura (V).....	41
Figura 12. Representação Simplificada do Neurônio Biológico .....	42
Figura13. Modelo de um Neurônio Artificial .....	43
Figura14. Ilustração de um MLP com duas camadas ocultas .....	46
Figura 15. Modelo descritivo de uma rede neural convolutiva. ....	48
Figura 16. Exemplo de operação de convolução. ....	49
Figura 17. Operação de <i>maxpooling</i> .....	50
Figura 18. Resultado da aplicação de <i>dropout</i> .....	51
Figura 19. Fluxograma da metodologia utilizada para a identificação e classificação de arritmias cardíacas.....	51
Figura 20. Diagrama em blocos mostrando o processo de extração e formação dos conjuntos de dados de cada classe e a formação de subconjuntos de 1600 amostras.....	55
Figura 21. Formação de um subconjunto pertencente a classe batimento prematuro atrial (a)....	55
Figura 22. Formação de um subconjunto pertencente a classe batimento de bloqueio do ramo esquerdo (L). ....	56
Figura 23. Formação de um subconjunto pertencente a classe batida de fuga nodal (j) .....	56
Figura 24. Processo de reescalonamento aplicado ao subconjunto. (a) Sinal de ECG original com o eixo vertical expresso em mV. (b) Sinal de ECG reescalonado para a faixa [0,255].....	58

Figura 25. (a) Sinal original de ECG com arritmia bloqueio do ramo esquerdo, mostrando a determinação dos valores máximo e mínimo de tensão (b) Sinal de ECG com os valores de tensão reescalados para a faixa [0,255] .....	59
Figura 26. (a) Sinal original de ECG com arritmia batimento prematuro atrial, mostrando a determinação dos valores máximo e mínimo de tensão (b) Sinal de ECG com os valores de tensão reescalados para a faixa [0,255].....	60
Figura 27. Processo de extração de um subconjunto a partir do sinal de ECG, reescalamento, conversão 1D-2D e geração de imagens 2D.....	61
Figura 28. (a) subconjunto do ECG contendo uma arritmia batimento prematuro nodal, a matriz com dimensões de 40x40 com valores reescalados; (b) imagem em escala de cinza de 8 bits obtida a partir dessa matriz. ....	62
Figura 29. Exemplo de aplicação da janela de captura. (a) janela com o complexo QRS contendo a arritmia bloqueio de ramo esquerdo centralizado; (b) janela com o complexo QRS contendo a arritmia bloqueio de ramo esquerdo à esquerda; (c) janela com o complexo QRS contendo a arritmia bloqueio de ramo esquerdo.....	63
Figura 30. Exemplo da técnica utilizada para aumento de dados em um sinal contendo a arritmia batimento prematuro nodal. São apresentados apenas 3 deslocamentos de 11 realizados. (a) janela com o complexo QRS contendo a arritmia batimento prematuro nodal à direita; (b) janela com o complexo QRS contendo a arritmia batimento prematuro nodal à esquerda; (c) janela com o complexo QRS contendo a arritmia batimento prematuro nodal centralizado .....	64
Figura 31. (a) arquitetura proposta CNN2D-1 contendo 3 blocos <i>conv-batch-relu</i> .(b) a arquitetura proposta CNN2D-2 contendo 5 blocos e (c) arquitetura proposta CNN2D-3 contendo 7 blocos .....	67
Figura 32. Fluxograma de como foi realizado o treinamento das três arquiteturas propostas .....	68
Figura 33. Trecho de um mapa de pontuação gerado pelo GradCAM.....	71
Figura 34. Exemplo de imagem aplicada à entrada da rede CNN2D, com seu respectivo mapa de calor gerado pelo método Grad-CAM.....	72
Figura 35. (a) Matriz de pontuação do método Grad. (b) Transformação da matriz em um vetor com 1600 coordenadas. (c) Correspondência entre as posições da matriz de peso e os trechos do subconjunto do sinal de ECG aplicado à entrada da rede CNN2D.....	73
Figura 36. Curva de convergência da acurácia ao longo do treinamento, para os conjuntos de validação e teste, utilizando-se o método SGDM com 3 blocos de extração de características sem aumento de dados .....	77

Figura 37. Matriz de confusão para o conjunto de teste, utilizando-se o método SGDM com 3 blocos de extração de características sem aumento de dados .....	77
Figura 38. Curva de convergência da acurácia ao longo do treinamento utilizando o aumento de dados, para os conjuntos de validação e teste, utilizando-se o método SGDM com 3 blocos de extração de características após o aumento de dados .....	78
Figura 39. Matriz de confusão para o conjunto de teste, utilizando-se o método SGDM com 3 blocos de extração de características após aumento de dados .....	79
Figura 40. Curva de convergência da acurácia ao longo do treinamento, para os conjuntos de validação e teste, utilizando-se o método SGDM com 5 blocos de extração de características sem o aumento de dados. ....	80
Figura 41. Matriz de confusão para o conjunto de teste, utilizando-se o método SGDM com 5 blocos de extração de características sem aumento de dados .....	80
Figura 42. Curva de convergência da acurácia para os conjuntos de validação e teste, utilizando-se o método SGDM .....	81
Figura 43. Matriz de confusão para o conjunto de teste, utilizando-se o método SGDM com 5 blocos de extração de características após aumento de dados .....	82
Figura 44. Curva de convergência da acurácia ao longo do treinamento, para os conjuntos de validação e teste, utilizando-se o método SGDM com 7 blocos de extração de características sem aumento de dados .....	83
Figura 45. Matriz de confusão para o conjunto de teste, utilizando-se o método SGDM com 7 blocos de extração de características sem aumento de dados .....	83
Figura 46. Curva de convergência da acurácia, utilizando-se o método SGDM com 7 blocos de extração de características após aumento de dados .....	84
Figura 47. Matriz de confusão para o conjunto de teste, utilizando-se o método SGDM com 7 blocos de extração de características após aumento de dados. ....	85
Figura 48. Subconjunto do sinal de ECG com destaque na cor verde, mostrando a presença de dois pontos contendo a arritmia batida de escape ventricular (E) .....	88
Figura 49. Subconjunto do sinal de ECG com destaque na cor verde, mostrando a presença da arritmia fusão de batimento ventricular (F).....	89
Figura 50. Subconjunto do sinal de ECG com destaque na cor verde, mostrando a presença de dois pontos contendo a arritmia bloqueio de ramo direito (R) .....	89
Figura 51. Subconjunto do sinal de ECG com destaque na cor verde, mostrando a presença da arritmia contração ventricular prematura (V).....	91
Figura 52. Subconjunto do sinal de ECG com destaque na cor verde, mostrando a presença da arritmia batida de fuga nodal (j) .....	91

## LISTA DE TABELAS

Tabela 1 – Síntese da revisão bibliográfica .....	29
Tabela 2. Exemplos de Registros da base MIT-BIH com número de eventos contabilizados .....	53
Tabela 3. As 17 classes de arritmia cardíaca na base MIT-BIH com número de eventos contabilizados. ....	53
Tabela 4. Classes de arritmias cardíacas da base de dados MIT-BIH a serem utilizadas neste trabalho. ....	54
Tabela 5. Classes selecionadas para o aumento de Dados.....	65
Tabela 6. Resultados obtidos para as três arquiteturas propostas utilizando o método de otimização ADAM .....	74
Tabela 7. Resultados obtidos para as três arquiteturas propostas utilizando o método de otimização SGDM .....	76
Tabela 8. Resultados das arquiteturas propostas após o aumento de dados utilizando o método de otimização SGDM. ....	86
Tabela 9. Desempenho do método proposto .....	86
Tabela 10. Valores de Coeficientes médios para cada classe de arritmia no conjunto de teste.....	87

## LISTA DE SIGLAS

ADAM: *Adaptive Moment Estimation*

AFL: *Flutter atrial*

AFIB: *Atrial fibrillation*

ANN: *Artificial neural network* APB:

Batimento prematuro atrial APC:

Contração prematura auricularAV:

Atrioventricular

CDVs: Doenças Cardiovasculares CNN:

*Convolutional Neural Network*

CNN1D: Rede convolutiva unidimensional

CNN2D: Rede convolutiva bidimensional

CWT: *Continuous Wavelet Transform* DCV:

Doença Cardiovascular

DT: *Decision tree classifier*

ECG: Eletrocardiograma FC:

*Fully Connected*

FN: Falso negativo

FP: Falso positivo

IVR: Ritmo idioventricular

ICA: Análise de Componentes Independentes

LBBBB: Batimento de bloqueio de ramo esquerdo LIME:

*Local Interpretable Model-agnostic Explanations*LSTM:

*Long Short-Term Memory Networks*

MIT-BIH: *Massachusetts Institute of Technology -Beth Israel Hospital*

ML: *Machine Learning*

MLP: *Multilayer Perceptron*

NSR: Ritmo sinusal normal

OMS: Organização Mundial da Saúde

PCDB: *PhysioNet Computing DataBase*

PVC: Contração Ventricular Prematura

ReLU: *Rectified Linear Unit*

RMSProp: *Root Mean Square Propagation*

RNA: Redes Neurais Artificiais

RBA: Ritmo de Batimento Atrial

RBBB: Bloqueio do ramo direito

SDHB: Bloqueio cardíaco de segundo grau

SGDM: *Stochastic Gradient Descent With Momentum*

SHAP: *SHapley Additive exPlanations*

SOBRAC: Sociedade Brasileira de Arritmias Cardíacas

SMOTE: *Synthetic Minority Oversampling Technique*

SVM: *Support Vector Machine*

SVTA: Taquiarritmia supraventricular

VEB: Batimento Ectópico Ventricular

VFW: Forma de Onda de Fibrilação Ventricular

VN: Verdadeiro negativo

VP: Verdadeiro positivo

VT: Taquicardia ventricular

WPW: *Wolff-Parkinson-White*

KNN: *K-nearest neighbors*

1D: Uma dimensão

2D: Duas dimensões

## SUMÁRIO

<b>1. INTRODUÇÃO .....</b>	<b>15</b>
1.1 OBJETIVO GERAL .....	18
1.2 OBJETIVO ESPECÍFICO.....	18
1.3 ORGANIZAÇÃO DO TRABALHO.....	18
<b>2. REVISÃO BIBLIOGRÁFICA. ....</b>	<b>20</b>
2.1 ANÁLISE DE ARTIGOS DO GRUPO 1.....	20
2.2 ANÁLISE DE ARTIGOS DO GRUPO 2.....	21
2.3 CONSIDERAÇÕES .....	28
<b>3. REFERENCIAL TEÓRICO .....</b>	<b>35</b>
3.1 O CORAÇÃO HUMANO .....	35
3.2 SINAL CARDÍACO.....	37
3.3 ELETROCARDIOGRAMA .....	38
3.4 ARRITMIAS CARDÍACAS .....	40
3.5 REDES NEURAIS ARTIFICIAIS (RNA).....	42
3.5.1 Funções de Ativações.....	45
3.5.2 <i>Perceptron</i> Multicamadas .....	45
3.5.3 Algoritmo <i>Backpropagation</i> .....	46
3.6 REDES NEURAIS CONVOLUCIONAIS (CNN).....	47
3.6.1 Convoluções.....	48
3.6.2 Camada de <i>Dropout</i> .....	50
<b>4. MATERIAIS E MÉTODOS. ....</b>	<b>51</b>
4.1 MATERIAIS.....	52
4.1.1 Reescalonamento do Sinal de ECG.....	57
4.1.2 Conversão do Vetor 1D em Matriz 2D .....	61
4.2 MÉTODOS .....	62
4.2.1 Aumento de Dados .....	62
4.2.2 Arquitetura da rede CNN proposta .....	65
4.2.3 Treinamento da rede.....	68
4.2.4 Métricas de Avaliação .....	68

4.2.5 Técnica de visualização de ativação .....	70
<b>5. RESULTADOS E DISCUSSÕES .....</b>	<b>74</b>
5.1 RESULTADOS COM MÉTODO DE OTIMIZAÇÃO ADAM.....	74
5.2 RESULTADOS COM MÉTODO DE OTIMIZAÇÃO SGDM.....	75
5.3 RESULTADOS DA TÉCNICA DE VISUALIZAÇÃO DE ATIVAÇÃO .....	87
<b>6. CONCLUSÃO.....</b>	<b>91</b>
<b>7. REFERÊNCIA.....</b>	<b>93</b>

## 1. INTRODUÇÃO

As doenças cardiovasculares (DCVs) são condições que afetam o coração e os vasos sanguíneos. Segundo a Organização Mundial da Saúde (OMS), as DCVs representam a principal causa de morte globalmente (OMS, 2020).

As DCVs englobam condições como a doença coronariana, que envolve o comprometimento do fluxo sanguíneo nas artérias coronárias (SWEIS e JIVAN, 2020). Além disso, a doença cerebrovascular é caracterizada por patologias que afetam os vasos sanguíneos do cérebro (FRANÇA, 2018). Outra condição relacionada é a doença cardíaca reumática, que é uma consequência da febre reumática aguda de longo prazo e resulta em inflamação nas válvulas cardíacas (THOMAS e KHETRAPAL, 2020).

No Brasil, essas doenças também têm impacto significativo na mortalidade. Segundo dados do Ministério da Saúde, em 2020, as DCVs foram responsáveis por mais de 30% das mortes em território brasileiro (Ministério da Saúde, 2020).

Segundo Machado *et al.* (2019) a detecção precoce de doenças cardiovasculares é fundamental para evitar complicações graves e permitir a intervenção terapêutica adequada. Dessa forma a detecção precoce desempenha um papel crucial na redução da morbidade e mortalidade relacionadas a essas condições de saúde.

De acordo com a *American Heart Association* (2021) "As arritmias cardíacas são perturbações no ritmo cardíaco, caracterizadas por batimentos acelerados, lentos ou irregulares. Podem ser assintomáticas ou provocar sintomas como palpitações, falta de ar, tontura e desmaios".

O ECG é o método diagnóstico mais utilizado para identificar e classificar arritmias cardíacas, infarto do miocárdio e doença arterial coronariana. Ele registra a atividade elétrica do coração fornecendo informações sobre o ritmo cardíaco e possíveis alterações na condução elétrica (Simões *et al.* 2018, p. 432) e (Beyer *et al.*, 2019, p. 235).

A identificação e classificação correta das arritmias desempenham um papel crucial no diagnóstico e no planejamento do tratamento. No entanto, é importante reconhecer que a interpretação do ECG é um processo complexo que requer conhecimento especializado e experiência clínica. Os médicos analisam várias características do ECG, como a morfologia das ondas, a duração dos intervalos e a regularidade do ritmo, para identificar padrões indicativos de arritmias (RAJKOMAR *et al.*, 2018).

Apesar da expertise dos médicos na interpretação de ECG, estudos têm demonstrado a existência de variabilidade interobservador na interpretação e diagnóstico

das arritmias cardíacas. Segundo Macfarlane *et al.* (2018), a concordância entre diferentes médicos na interpretação de ECG pode variar, especialmente em casos de arritmias complexas ou quando há sobreposição de características.

Tais fatos, além do volume de realização de ECGs nos ambientes ambulatoriais e hospitalares e a respectiva demanda de laudos desses exames têm motivado o desenvolvimento de técnicas computacionais de apoio à decisão clínica. Sistemas de apoio diagnóstico baseados em algoritmos de processamento de sinais e imagens têm demonstrado capacidade de auxiliar os médicos na detecção e classificação de arritmias com precisão e consistência (SHAH *et al.*, 2020).

O uso de técnicas computacionais, como o processamento de sinais de ECG, tem se mostrado eficaz nesse contexto. A análise computacional de ECG permite extrair informações relevantes sobre a atividade elétrica do coração e identificar padrões característicos das diferentes arritmias. Essas técnicas têm sido aplicadas com sucesso no desenvolvimento de algoritmos de detecção e classificação de arritmias cardíacas.

Na revisão da literatura realizada neste trabalho, foram identificados dois grupos de trabalhos desenvolvidos com o objetivo de realizar a detecção automática de arritmias. O primeiro grupo compreende trabalhos com ferramentas clássicas de aprendizado de máquina, enquanto o segundo grupo compreende trabalhos que utilizam técnicas de aprendizado profundo, como redes convolucionais e redes recorrentes.

Referente ao primeiro grupo, analisamos o trabalho de Singh *et al.*(2019) que realizou a detecção de arritmias e classificação do sinal de ECG em 5 classes utilizando as seguintes técnicas clássicas de aprendizado de máquina: Classificadores utilizando máquinas de vetores de suporte (do inglês *Support Vector Machine* – SVM) e floresta aleatória (do inglês *Random Forest*).

Referente ao segundo grupo citado, analisamos os trabalhos de Yldirim *et al.* (2018), Abdullah e Alani (2020) e Petmezas *et al.* (2021) que utilizaram redes neurais convolucionais unidimensional (CNNs), e tinham como sinal de entrada vetores com amostras retiradas do ECG. Ainda no segundo grupo, foram analisados os trabalhos dos autores Rohmantri e Surantha (2020) e Zeng *et al.* (2020) que utilizaram redes convolucionais bidimensionais e na sua entrada imagens capturadas do ECG. A semelhança destes dois últimos autores, segue Ajjeey *et al.* (2022) e Ullah *et al.* (2020), entretanto as imagens utilizadas são retiradas de espectrogramas.

Foi observado em alguns trabalhos o uso de redes híbridas, que consistem em uma composição em série de redes convolucionais e redes recorrentes. Nesse contexto, as redes convolucionais são empregadas para extrair características relevantes dos dados. Essas características extraídas são então submetidas a uma redução de dimensionalidade, por meio de camadas de amostragem, e são utilizadas como entrada para uma rede recorrente (RAJPURKAR *et al.*, 2017).

A base de dados selecionada para a extração dos sinais de ECG, foi a “MIT-BIH *Arrhythmia Database*” coletada pelo Laboratório de Arritmias do Hospital Beth Israel em parceria com o Instituto de Tecnologia de Massachusetts. Essa base está disponível no banco de dados da PhysioNet (PHYSIONET, 2021). Os sinais de ECG, foram obtidos a partir de 47 pacientes e amostrados a uma taxa de 360 Hz. A base é composta por 48 registros anotados, com duração aproximada de 30 minutos cada um. Os registros contêm dados de 17 tipos de arritmias.

Nesta pesquisa estão sendo utilizadas 13 classes do referido conjunto de dados para a realização de classificação (12 tipos de arritmias cardíacas e a classe de ECG normal). Também está sendo proposto um novo método de aumento de dados, visto que entre os 12 tipos de arritmias selecionadas, 6 possuem quantidades de amostras pequenas em relação as outras. Visando minimizar esse desbalanceamento de amostras entre as classes, fez-se necessário o uso de técnicas de aumento de dados.

Em grande parte da literatura atual, são comuns a utilização de imagens provenientes de trechos de ECG. Em nosso trabalho de pesquisa, propomos um novo modelo de Rede Convolutiva Bidimensional (CNN2D). Neste modelo, empregamos dados unidimensionais obtidos a partir de registros que foram inicialmente transformados em sinais bidimensionais. Posteriormente, por meio de técnicas avançadas de processamento de imagens, esses sinais foram convertidos em imagens de intensidade com dimensões de 40x40 pixels.

Em resumo, esta pesquisa propõe um método que utiliza uma imagem de intensidade obtida pela transformação do ECG em sinais 2D para classificação de arritmias cardíacas. A imagem de intensidade é obtida convertendo os valores da amostra de ECG para valores em tons de cinza de 8 bits (variando entre 0 e 255). A classificação é realizada por uma rede neural convolucional 2D, são apresentadas três propostas de arquiteturas de redes convolucionais (CNN2D-1, CNN2D-2 e CNN2D-3) e seus respectivos desempenhos. Para permitir a comparação com outros estudos previamente

publicados na literatura, foi utilizada a base de dados pública. Tendo em vista o banco de dados ser desbalanceado, uma nova técnica de aumento de dados é proposta. Nessa técnica, novas amostras são obtidas por meio de uma janela deslizante ao redor do segmento do sinal de ECG que contém a arritmia. Para justificar a decisão do classificador, a técnica Grad-CAM é empregada, gerando de forma visual o destaque do trecho do ECG que mais contribuiu com a decisão da rede em classificar o sinal em uma arritmia cardíaca específica ou como sinal normal.

### **1.1 Objetivo Geral**

Implementar um método para detecção e classificação de arritmias cardíacas em sinais de ECG, utilizando modelos neurais profundos e fornecer explicações sobre a decisão do modelo proposto.

### **1.2 Objetivos Específicos**

1. Avaliar se há aumento de desempenho nos modelos neurais profundos propostos para detecção e classificação de arritmias em sinais de ECG ao utilizarmos dados bidimensionais;
2. Avaliar se há aumento de desempenho nos modelos neurais profundos propostos para detecção e classificação de arritmias em sinais de ECG com a utilização de uma nova técnica de aumento de dados em sinais de ECG;
3. Avaliar o potencial de técnicas de visualização de ativação para identificação de regiões com arritmia em sinais de ECG.

### **1.3 Organização do Trabalho**

Este trabalho está organizado conforme a divisão descrita a seguir:

- Capítulo 1: Introdução;
- Capítulo 2: Revisão Bibliográfica;
- Capítulo 3: Referencial Teórico;
- Capítulo 4: Materiais e Método;
- Capítulo 5: Resultados e discussão;
- Capítulo 6: Conclusão
- Capítulo 7: Referências Bibliográficas

O Capítulo 1 contextualiza o tema que será trabalhado, doenças cardiovasculares, arritmias cardíacas, suas consequências, a importância da sua detecção de forma precoce e quais os objetivos gerais e específicos deste trabalho.

O Capítulo 2 apresenta trabalhos que foram publicados na literatura na área de detecção e classificação de exames de ECG por meio de redes neurais convolucionais, bem como o uso de técnicas de transformações desses exames em imagens. por meio de técnicas de aprendizado de máquina e redes neurais convolucionais.

O Capítulo 3 apresenta os principais fundamentos teóricos utilizados no desenvolvimento deste trabalho: o coração humano e sua estrutura, os sinais elétricos gerados pelo coração, eletrocardiogramas, as arritmias cardíacas, redes neurais artificiais, redes neurais convolucionais.

No Capítulo 4 são apresentados os materiais utilizados nessa proposta de projeto, as características do banco de dados, como foram extraídas e convertidas as informações, o processo de geração de imagens, a técnica de aumento de dados, arquitetura proposta, fluxo do treinamento, validação e testes, e as métricas utilizadas.

No Capítulo 5 são apresentados os resultados que foram alcançados e discutido esses resultados.

No Capítulo 6 é apresentada a conclusão deste trabalho e por fim, no capítulo 7 são apresentadas as referências utilizadas.

## 2. REVISÃO BIBLIOGRÁFICA

Neste capítulo é apresentada a revisão bibliográfica realizada sobre o tema “detecção e classificação de arritmias em sinais de ECG”. Um conjunto de quatorze artigos entre os anos de 2018 e 2024 foi selecionado a partir da busca nas seguintes bases de dados literárias: *IEEE xplora* e *Web of Science*. Para fins de comparação com os resultados de nossa proposta de trabalho, as buscas se restringiram a artigos que utilizaram a base de dados pública MIT-BIH ECG. A análise dos artigos visou a identificação dos métodos (pré-processamento e classificação) e dos resultados obtidos. Durante a revisão bibliográfica os artigos foram divididos em dois grupos. O grupo 1 refere-se a um trabalho que utiliza técnicas clássicas de aprendizado de máquina. O grupo 2 refere-se a trabalhos que utilizam técnicas de aprendizado de profundo. A seguir será feita uma análise dos trabalhos pertencentes a esses dois grupos. No final dessa revisão é montada uma Tabela sumariando os pontos destacados e após isso são levantadas algumas considerações que motivaram o tema desta pesquisa.

### 2.1 Análise do trabalho do grupo 1

#### 2.1.1 *A Machine Learning based Comparative Analysis with MIT-BIH ECG Data (SINGH et al., 2019)*

Nesse trabalho os autores realizaram a classificação de sinais de ECG utilizando a base de dados MIT-BIH, adotaram uma classificação de cinco classes: normal, batida compassada, batida de bloqueio de ramo direito, batimento de bloqueio de ramo esquerdo e batimento ventricular prematuro. Foram utilizadas 25000 amostras de sinais divididas entre as cinco classes (5000 amostras/classe). Na etapa de pré-processamento, foi realizada uma filtragem de ruídos (através de um filtro de média) com um kernel de tamanho igual a 300 amostras. O algoritmo utilizado para a detecção do complexo QRS, através da identificação do pico R, foi o de Pan-Tompkins (1985). Os sinais utilizados são centralizados em torno do complexo QRS e são formados por 300 amostras, 149 amostras à esquerda e 150 amostras na direita do pico R. A partir dos sinais gerados, os autores calcularam a área sobre a curva gerada pelo complexo QRS através da regra *3/8 de Simpson*, obtendo valores individuais característicos para cada complexo. Para a classificação das arritmias, os autores utilizaram cinco métodos de aprendizado supervisionado: máquina de vetores de suporte (SVM - do inglês *Support Vector Machine*), classificador de árvore de decisão (do inglês *Decision tree classifier*),

classificador de floresta aleatória (do inglês *Random forest classifier*), classificador de Naïve Bayes e redes neurais artificiais (ANN – do inglês *Artificial neural network*). Utilizando a acurácia como métrica de avaliação, a ANN teve o melhor desempenho (99,59%), seguida do classificador de floresta aleatória (99,08%) e da SVM(98,97%).

## **2.2 Análise dos trabalhos do grupo 2**

### **2.2.1 *CNN-LSTM Based Model for ECG Arrhythmias and Myocardial Infarction Classification.* (ABDULLAH e ALANI, 2020)**

Nesse trabalho os autores desenvolveram um classificador de sinais de ECG objetivando a detecção de infarto do miocárdio. Foram utilizadas duas bases de dados, MIT-BIH para o classificador e PTP Physionet (PHYSIONET, 2019) para o detector de infarto no miocárdio. A proposta foi de um modelo de rede híbrida CNN-LSTM. A rede CNN foi utilizada em duas etapas, a primeira etapa refere-se à extração de características, para isso foram combinadas as seguintes camadas: duas *Conv-Batch-Relu*, uma *Conv-Batch-Relu-Dropout* e uma *Conv-Batch-Softmax-Dropout*. A segunda etapa, refere-se à classificação, constituída de uma *softmax* com duas saídas. A rede LSTM é constituída por duas LSTM, uma de *dropout* e uma *softmax* na saída. Utilizando a acurácia como métrica de avaliação, os classificadores de sinais de ECG apresentaram os seguintes resultados: 94,7% para a CNN e de 98,66% para CNN-LSTM.

### **2.2.2 *Automated atrial fibrillation detection using a hybrid CNN-LSTM Network on Imbalanced ECG Datasets* (PETMEZAS *et al.*, 2020)**

Nesse trabalho os autores desenvolveram um detector de fibrilação atrial utilizando a base de dados MIT-BIH, adotaram quatro classes: fibrilação atrial, vibração atrial, ritmo do nodo atrioventricular e normal. Na etapa de pré-processamento os autores utilizaram o filtro *Butterworth* passa alta de 7ª ordem e um *Butterworth* passa baixa de 6ª ordem, para remoção do ruído da linha de base e do ruído da rede de alimentação, respectivamente. As características extraídas dos batimentos foram em torno do complexo QRS, para identificá-lo foi utilizada a transformada discreta *wavelet* de 4 níveis. Após a detecção do pico R, uma amostra do sinal é colhida a partir de uma janela em torno dele com 250ms à esquerda e 500ms à direita. A proposta dos autores foi utilizar uma rede híbrida CNN-LSTM. A rede CNN-1D, é responsável pela extração de

características através das camadas *Conv-Batch-Relu*. Essas características são repassadas à rede LSTM. Os autores adotaram como métricas: a sensibilidade e especificidade. Os melhores resultados foram 97,87% de sensibilidade e 99,29% de especificidade.

### **2.2.3 *Arrhythmia Classification using 2D Convolutional Neural Network (ROHMANTRI e SURANTHA, 2020).***

Nesse trabalho os autores realizaram a classificação de sinais de ECG utilizando a base de dados MIT-BIH, estabeleceram três etapas: a primeira utilizando apenas 2 classes: sinal normal e sinal com arritmia, a segunda etapa contendo 7 classes e terceira contendo 8. As arritmias trabalhadas foram: contração atrial prematura, contração ventricular prematura, batimento prematuro, fusão de batimento normal e ritmado, fusão de batimento ventricular e normal, bloqueio do ramo esquerdo e bloqueio do ramo direito. A rede proposta foi uma CNN2D. Os autores realizaram testes usando na entrada da rede imagens nas dimensões 16x16, 32x32 e 64x64 pixels. Durante a fase de treinamento, os autores utilizaram validação cruzada com 10 pastas. Os resultados obtidos foram medidos através da acurácia: 98,91% para duas classes, 98,10% para 7 classes e 98,45% para 8 classes. Os autores concluíram que utilizar imagens menores, apesar de reduzir o processamento, não aumentam o valor da acurácia, ou seja, os melhores resultados são obtidos quando as imagens possuem dimensões iguais ou maiores que 128x128 pixels.

### **2.2.4 *An Automatic Diagnosis of Arrhythmias Using a Combination of CNN and LSTM Technology (ZHENG et al.,2020)***

Nesse trabalho os autores propõem o diagnóstico automático de arritmia cardíaca. A base de dados utilizada foi a MIT-BIH. A proposta foi de classificar os ECG em sete classes: bloqueio de ramo esquerdo, bloqueio de ramo direito, batimento atril prematuro, contração prematura ventricular, batimento estimulado, batimento de escape ventricular e onda de vibração ventricular e uma classe de sinal normal: ritmo sinusal normal. Os sinais foram extraídos a partir do complexo QRS, tendo como marcação o pico da onda R, isso foi feito com base nos marcadores de pico da onda R existentes no banco de dados. Em seguida foi estabelecida uma janela com 92 amostras a direita e 92 amostras a esquerda dos picos da onda R, essas amostras foram utilizadas para gerar e as imagens 2D de dimensão 192x128 pixels. Por haver um desbalanço entre as quantidades de

imagens geradas, foi necessário realizar o aumento de dados, os autores realizaram a técnica de recorte, onde uma imagem original é subdividida em outras imagens, o fato se deu pela formação de quadrados onde os autores determinaram as coordenadas dos quatro vértices, o primeiro vértice foi na origem (0,0), o segundo (0,96), terceiro (96,0) e quarto (96,96), resultando em uma nova imagem 96x96 pixels, essa etapa foi repetida mais 8 vezes alterando as coordenadas, após isso, as imagens são redimensionadas para o tamanho padrão adotado (192x128 pixels), isto permitiu aumentar de 107.620 para 581.148 imagens de ECG. Os autores utilizaram a combinação de duas redes, CNN2D – LSTM, composta por 13 camadas. As 9 primeiras camadas repetem a formação Conv2D – Conv2D – Pool, em seguida uma camada LSTM, duas camadas *fully-connected* de saída. Os resultados obtidos foram: acurácia de 99,01%, a especificidade de 99,57% e a sensibilidade de 97,67%.

### ***2.2.5 Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats (OH et al., 2018)***

Nesse trabalho os autores desenvolveram um sistema automatizado de diagnóstico para arritmia cardíaca. A base de dados utilizada foi a MIT-BIH, foram selecionadas cinco classes, uma normal: ritmo sinusal normal e quatro de arritmia: bloqueio de ramo esquerdo, bloqueio de ramo direito, batimentos prematuros atriais e contração ventricular prematura. Cada sinal unidimensional foi extraído a partir do complexo QRS, tendo como marcação central o pico da onda R sendo coletados 99 registros a esquerda e 160 registros a direita. A rede proposta é um modelo híbrido de aprendizado profundo que combina CNN com LSTM. A rede possui 10 camadas, sendo as 6 primeiras compostas pela repetição da sequência *Conv1D – Pool*, uma camada LSTM e finalizando com 3 camadas de *fully-connected*. Os autores utilizaram durante o treinamento, a validação cruzada com dez pastas. Os resultados obtidos foram: acurácia de 98,42%, precisão de 98,10%, sensibilidade de 97,50% e especificidade de 98,70%.

### ***2.2.6 Arrhythmia detection using deep convolutional neural network with long duration ECG signals (YLDIRIM et al., 2018)***

Nesse trabalho os autores desenvolveram um detector de arritmia cardíaca utilizando sinais de ECG. A base de dados utilizada foi a MIT-BIH, os autores realizaram testes com 13, 15 e 17 classes, foram elas: ritmo sinusal normal, batimento atrial

premature, vibração atrial, fibrilação atrial, taquiarritmia supraventricular, pré-excitação, contração ventricular prematura, bigeminia ventricular, trigeminia ventricular, taquicardia ventricular, ritmo idioventricular, vibração ventricular, fusão do batimento ventricular e normal, batida do ramo esquerdo, batida do ramo direito, bloqueio cardíaco de segundo grau e ritmo de marcapasso. Cada sinal obtido continha 3600 amostras, essas amostras foram extraídas de trechos dos sinais de ECG com duração de 10 segundos. Não foram utilizados filtros e não foi necessário localizar os complexos QRS. A rede proposta pelos autores foi uma CNN1D, composta por 16 camadas. A seção de extração de características é constituída por blocos formados pelas seguintes camadas: *Conv-Batch-Maxpool*. Os valores de acurácia para a classificação de 13, 15 e 17 classes, respectivamente, foram 95,2%, 92,51% e 91,33%.

### ***2.2.7 Scalogram Based Heart Disease Classification using Hybrid CNN-Naive Bayes Classifier (AJJEY et al., 2022)***

Neste trabalho os autores propuseram um classificador de doenças cardíacas baseada em espectrogramas. A base de dados utilizada foi a MIT-BIH e foram utilizadas ao todo 106 segmentos de ECG, totalizando um conjunto de 3 classes: Ritmo Sinusal, insuficiência cardíaca congestiva e arritmia anormal. Cada sinal unidimensional foi composto de 1000 amostras e foram convertidos primeiramente em um gráfico visual no domínio da frequência através da Transformada *Wavelet* Contínua (CWT) e posteriormente gerada uma imagem colorida do gráfico. De acordo com os autores, a principal vantagem da representação em imagens 2D dos espectrogramas, em comparação com os sinais unidimensionais, reside na capacidade de preservar detalhes e batimentos, evitando perdas de informações cruciais. A arquitetura adotada envolveu uma rede híbrida, combinando uma CNN2D com um classificador Naive Bayes. A CNN2D desempenhou o papel de extrair características significativas das imagens dos espectrogramas, enquanto o classificador Naive Bayes foi responsável pela classificação final. Como métrica de avaliação, a acurácia foi empregada, alcançando o valor de 98,76%.

### ***2.2.8 Multiclass Heartbeat Classification using ECG Signals and Convolutional Neural Networks (QURESHI et al., 2022)***

Neste trabalho os autores propuseram um classificador de batimento cardíaco

multiclasse. A base de dados utilizada foi a MIT-BIH e foram utilizadas ao todo 10 classes: Batida Normal (N), Batida de bloqueio de ramo direito (R), Batida compassada (/), Batimento de bloqueio de ramo esquerdo (L), Contração ventricular prematura (V), Batimento prematuro atrial(A), Fusão de batimento compassado e normal (f), Batimento prematuro nodal (J), Batida inclassificável (Q) e Batida de fuga atrial (e). Cada sinal unidimensional foi extraído a partir do complexo QRS, tendo como marcação central o pico da onda R, com duração de 0,65s (0,25s antes do pico R e 0,4s após o pico R). Com o objetivo de melhorar o processo de aprendizado, os autores utilizaram a técnica de aumento de dados SMOTE (*Synthetic Minority Oversampling Technique*) que é baseado em um classificador KNN (*K-nearest neighbors*), o KNN classifica cada amostra de um conjunto de dados avaliando sua distância em relação aos vizinhos mais próximos. A arquitetura utilizada no trabalho consiste em uma rede CNN1D, combinada com o método de otimização ADAM. Os autores observaram que o ADAM tendeu a convergir mais rapidamente para o mínimo global, eliminando a necessidade de ajustes minuciosos na taxa de aprendizado. Como métrica de avaliação, a acurácia foi empregada, alcançando o valor de 95,7%.

### **2.2.9 ECG Beat Classification using CNN (Deepak e Vijaykumar, 2022)**

Neste trabalho os autores propuseram um classificador de batimento cardíaco. Os dados utilizados foram extraídos da base MIT-BIH e foram utilizadas cinco classes de batimentos cardíacos: Batida normal (N), batimento ectópico supraventricular (S), batimento ectópico ventricular (V), fusão do batimento ventricular (F), batimento cardíaco não diagnosticado (U) e diferentes tipos de batimentos cardíacos (Q). A extração dos dados foi realizada a partir de segmentos de ECG, onde foram identificados os complexos QRS através de marcações já existentes. Cada segmento tem um comprimento total de 10s que equivale a 3600 amostras. A arquitetura proposta pelos autores trata-se de uma adaptação do modelo U-Net que é uma rede Convolutacional (CNN1D). A métrica proposta para avaliação foi a acurácia, sendo obtido um valor de 90,4%.

### **2.2.10 Arrhythmia detection model using modified DenseNet for comprehensible Grad-CAM visualization (KIM et al., 2021)**

Nesse trabalho os autores propuseram um detector de arritmia cardíaca automático com indicação visual do trecho contendo a arritmia. A base de dados utilizada nesse trabalho totalizou 52000 segmentos de ECG de pacientes coletados em hospitais

universitários na Coreia do Sul. Trata-se de um detector de classificação binária (sinais normais e sinais anormais). Para a formação da classe de sinais anormais os autores utilizaram ao todo 12 tipos de arritmias reunidas formando um único conjunto. Os 12 tipos de arritmias cardíacas são: fibrilação atrial, APC, flutter atrial (FA), dissociação AV, ritmo de escape juncional, arritmia sinusal, bradicardia sinusal, taquicardia, taquicardia supraventricular, taquicardia ventricular prematura, contração e taquicardia de QRS largo. No pré-processamento com o objetivo de padronizar todos os sinais unidimensionais, foram realizadas interpolações lineares tornando as amostras sinais na faixa de 200Hz e cada sinal foi formado a partir de trechos de ECG com tamanho de 10s de duração. A arquitetura proposta foi o modelo DenseNet, que trabalha com redes convolutivas densas que ao contrário das redes neurais convolucionais tradicionais (CNNs) a DenseNet estabelece conexões diretas entre todas as camadas garantindo que cada camada receba não apenas os mapas de características da camada anterior, mas também os mapas de características de todas as camadas anteriores. Na última etapa, com o objetivo de mostrar visualmente o resultado da classificação da rede, foi utilizada a técnica GradCAM, onde os trechos em que a rede classificou conter a arritmia foram destacadas através da cor cinza. A métrica utilizada para avaliação de desempenho foi a acurácia, e neste trabalho foi de 98,71%.

### ***2.2.11 Classification of Arrhythmia by Using Deep Learning with 2-D ECG Spectral Image Representation (ULLAH et al., 2020)***

Nesse trabalho os autores propuseram um classificador de arritmias cardíacas utilizando imagens de espectrogramas obtidos a partir de trechos de ECG. A base utilizada foi a MIT-BIH e foram utilizadas ao todo 8 classes, sendo 7 de arritmias cardíacas e uma de batimentos normais. Os tipos de arritmias cardíacas utilizadas foram: Ritmo Sinusal Normal (NOR), Forma de Onda de Fibrilação Ventricular (VFW), Contração Ventricular Prematura (PVC), Batimento Ectópico Ventricular (VEB), Bloqueio do Ramo Direito (RBB), Bloqueio do Ramo Esquerdo (LBB), Ritmo de Batimento Atrial (RBA) e Contração Atrial Prematura (APC). Inicialmente foram extraídos sinais unidimensionais do ECG e então transformados em espectrogramas 2D por meio de transformada de Fourier, após isso são capturadas imagens dos espectrogramas no formato 256x256 pixels. Os autores aplicaram técnica de aumento de dados com o objetivo de melhorar o aprendizado, a técnica utilizada foi a de recorte, onde as imagens dos espectrogramas originais (255x256 pixels) são recortadas, originando

pedaços menores de espectrogramas de tamanho 200x200 pixels. Por a classe de batimentos normais ter uma grande quantidade de amostras, essa técnica foi utilizada apenas nas classes de arritmias. Os autores propuseram como arquitetura as redes convolutivas CNN2D, uma vez que foram analisadas imagens 2D dos espectrogramas e isso é uma especialidade delas. As métricas utilizadas com seus respectivos valores foram: sensibilidade média de 97,91%, especificidade de 99,61%, acurácia média de 99,11% e precisão de 98,59%.

#### **2.2.12 Arrhythmia Detection from ECG Signals using CNN Model (KUMAR *et al.*, 2023).**

Neste estudo, os autores propuseram um classificador de arritmias cardíacas em sinais de ECG, utilizando a base de dados MIT-BIH. Foram empregadas cinco classes de batimentos cardíacos: Batimentos Não Ectópicos (N), Ectópicos Supraventriculares (S), Batimentos Ectópicos Ventriculares (V), Batimentos de Fusão (F) e Batimentos Desconhecidos (Q). Os dados de entrada consistiam em sinais unidimensionais, sem a aplicação de técnicas de aumento de dados. Cada sinal extraído era composto por 360 amostras, centradas no pico R. Para normalizar os valores das amostras, foi utilizada a técnica de Normalização Z-score, indicada para dados com distribuição normal ou aproximadamente normal. O Z-score calcula a diferença entre um valor e a média do conjunto de dados, dividida pelo desvio padrão, destacando valores discrepantes e tornando os dados comparáveis (Walpole, 2015). A proposta dos autores foi o uso de uma rede CNN unidimensional. As métricas de avaliação com seus respectivos valores foram: Acurácia de 98.46%, Sensibilidade de 97.75%, Especificidade de 98.59%, Precisão de 92.32% e Pontuação F1 de 96.46%.

#### **2.2.13 ECG-Signals-based Heartbeat Classification: A Comparative Study of Artificial Neural Network and Support Vector Machine Classifiers (UGWU *et al.*, 2024).**

Nesse estudo, os autores propuseram um classificador de arritmias cardíacas em sinais de ECG. A base de dados utilizada foi a MIT-BIH, foram utilizadas cinco classes de batimentos cardíacos: Batimentos Não Ectópicos (N), Ectópicos Supraventriculares (S), Batimentos Ectópicos Ventriculares (V), Batimentos de Fusão (F) e Batimentos Desconhecidos (Q). Eles utilizaram como sinal de entrada dados unidimensionais, esses

sinais foram inicialmente pré-processados e extraídos utilizando a Transformada Wavelet Discreta (DWT), seguido pela aplicação da Análise de Componentes Independentes (ICA) para a obtenção de recursos de características avançadas. Não foram aplicadas técnicas de aumento de dados. As amostras dos sinais foram identificadas centradas nos picos R, com cada batimento padronizado para um comprimento fixo de 128 Hz e normalizado. O objetivo dos autores foi realizar uma análise comparativa de desempenho entre redes neurais artificiais (RNAs) e máquinas de vetores de suporte (SVM). As métricas utilizadas de avaliação e os resultados obtidos foram: para o modelo RNA uma acurácia de 98,40%, enquanto o modelo SVM alcançou 98,70%.

A Tabela 1 mostra um resumo dos artigos analisados sobre métodos automáticos para diagnóstico de arritmia cardíacas. As seguintes características desses artigos são apresentadas: banco de dados utilizado, sinal de entrada, classificador, método de extração de dados, técnica de aumento de dados, número de classes e resultados.

### 2.3 Considerações

A maioria dos artigos da Tabela 1 usa o banco de dados MIT-BIH. Nesta base de dados, os sinais de ECG, obtidos de 47 pacientes, são amostrados a uma taxa de 360 Hz. São 48 registros anotados, com duração aproximada de 30 minutos de pulsação. 17 classes estão presentes no banco de dados: ritmo sinusal normal (NSR), ritmo de marcapasso (PR) e as seguintes 15 disfunções cardíacas: batimento prematuro atrial (APB), *flutter* atrial (AFL), fibrilação atrial (AFIB), taquiarritmia supraventricular (SVTA), pré-excitação (WPW), contração ventricular prematura (PVC), bigeminismo ventricular (B), trigeminismo ventricular (T), taquicardia ventricular (VT), ritmo idioventricular (IVR), flutter ventricular (VFL), fusão de ventrículos e batimento normal (F), batimento de bloqueio de ramo esquerdo (LBBBB), batimento de bloqueio de ramo direito (RBBB) e bloqueio cardíaco de segundo grau (SDHB).

Os seguintes tipos de sinais de entrada são usados pelos métodos mostrados na Tabela 1: vetores, imagens binárias 2D e espectrogramas 2

Tabela 1 – Síntese da revisão bibliográfica

	<b>Artigo</b>	<b>Banco de dados</b>	<b>Sinal de entrada/classificador</b>	<b>Extração de dados/ Aumento de dados</b>	<b>Classes</b>	<b>Resultados (%)</b>
1	ZHENG <i>et al.</i> (2020)	MIT-BIH (107.620 batimentos cardíacos)	Imagem binária -2D/ CNN2D-LSTM	Batimentos de ECG (92 amostras antes e depois do pico da onda R) /Sim. (9 métodos diferentes de recorte de imagem)	8	Acc:99,01 Sens:97,67 Espec:99,57
2	YILDIRIM <i>et al.</i> (2018)	MIT-BIH (3600 Segmentos de sinais de ECG)	Vetor/ CNN1D	Segmentos de ECG de longa duração (10s) /Não	13 15 17	Acc:95,20 Acc:92,51 Acc: 91,33
3	ROHMANTRI <i>et al.</i> (2020)	MIT-BIH (108,819 batimentos cardíacos)	Imagem binária -2D/ CNN2D	Batimentos de ECG (não detalha como as amostras são extraídas) /Não	2 7 8	Acc: 98,91 Acc: 98,10 Acc: 98,45
4	OH <i>et al.</i> (2018)	MIT-BIH (16499 segmentos de sinal de ECG)	Vetor/ CNN1D/LSTM	Segmentos de ECG com 1000 amostras cada/ Não	5	Acc: 98,42
5	AJJEY <i>et al.</i> (2022)	MIT-BIH (162 segmentos de sinal de ECG)	Vetor/ transformada wavelet contínua + CNN2D+Nayve Bayes	Segmentos de ECG com 1000 amostras cada/ Não	3	Acc: 98,76

	<b>Artigo</b>	<b>Banco de dados</b>	<b>Sinal de entrada/classificador</b>	<b>Extração de dados/ Aumento de dados</b>	<b>Classes</b>	<b>Resultados (%)</b>
6	QURESHI <i>et al.</i> (2022)	MIT-BIH (103.214 batimentos cardíacos)	Vetor/CNN1D	Batimentos cardíacos de ECG com 0,65s (0,25s antes do pico R e 0,4s após o pico R)/ Sim ( <i>Synthetic Minority Oversampling Technique - SMOTE</i> )	10	Acc: 95,70
7	DEEPAK <i>et al.</i> (2022)	MIT-BIH (Sem menção ao número de segmentos)	Vetor/CNN1D	Segmentos de ECG com 3600 amostras/ Não	5	Acc: 90,40
8	PETMEZAS <i>et al.</i> (2020)	Fibrilação atrial MIT-BIH (970.009 batimentos cardíacos)	Vetor/ CNN1D-LSTM	Batimentos cardíacos de ECG com 0,75s (0,25s antes do pico R e 0,5s após o pico R)/Sim (perda focal)	4	Acc: 97,87
9	SINGH <i>et al.</i> (2022)	MIT-BIH (108.819 batimentos cardíacos)	Vetor/ máquina de vetores de suporte, árvore de decisão, floresta aleatória, Naive-Bayes, rede neural artificial	Batimentos cardíacos de ECG com 300 pontos de amostragem (0,42s antes e depois do pico R)/ Não	5	Acc SVM: 98,97 DT: 97,61 RF: 99,08 NB:83,61 ANN: 99,59

	<b>Artigo</b>	<b>Banco de dados</b>	<b>Sinal de entrada/classificador</b>	<b>Extração de dados/ Aumento de dados</b>	<b>Classes</b>	<b>Resultados (%)</b>
10	KIM <i>et al.</i> (2021)	Banco de Dados de Hospitais da Coréia do Sul / 52.043 segmentos	Vetor/CNN1D	Segmentos de ECG de 10s/ Não	2	Acc: 98,71
11	ABDULLA <i>et al.</i> (2020)	MIT-BIH	Vetor/CNN1D-LSTM	Batimentos cardíacos de ECG com 130 amostras (0,18s antes e depois do pico R) / Não	2	Acc: 98,66
12	ULLAH <i>et al.</i> (2020)	MIT-BIH (108.819 batimentos cardíacos)	Espectrogramas 2D/ CNN2D	Sinais unidimensionais de série temporal de ECG são transformados em espectrogramas 2-D por meio de transformada de Fourier de curta duração/ Sim (método de corte)	8	Acc: 99,11
13	KUMAR <i>et al.</i> , (2023)	MIT-BIH	Vetor/CNN1D	Batimentos cardíacos de ECG com 360 amostras, Normalização Z-score/ Não	5	Acc: 98,46
14	UGWU <i>et al.</i> , (2024)	MIT-BIH	Vetor/ rede neural artificial, máquina de vetores de suporte	Sinais unidimensionais extraídos através da Transformada Wavelet Discreta seguido da Análise de Componentes Independentes/ Não	5	ANN Acc: 98,40 SVM Acc: 98,70

Acc: acurácia, Sens: sensibilidade, Espec: especificidade.

Para construir as entradas do vetor, um complexo QRS ou um segmento de ECG de longa duração são extraídos. Na literatura científica, a análise do complexo QRS é substancialmente mais popular do que a análise de fragmentos de sinal de ECG de longa duração. Na Tabela 1, 9 estudos empregam análise do complexo QRS, enquanto 3 empregam análise de fragmentos de ECG de longa duração. De acordo com Yildirim *et al.*, os métodos que utilizam a análise do complexo QRS podem ser propensos a erros e podem não alcançar desempenho diagnóstico satisfatório devido à alta variabilidade batimento a batimento dessas características entre os indivíduos. Na Tabela 1, para análise do complexo QRS, as amostras são extraídas em torno dos picos R, fixando várias amostras antes e depois do pico R (65, 150), ou um intervalo de tempo antes e depois do pico R (0,25s antes e 0,4s depois, 0,42s antes e depois, 0,18s antes e depois). Para análise de segmentos longos, o intervalo de amostragem é fixado em termos de número de amostras (1000, 3600) ou em termos de tempo (10s). Alguns trabalhos que usam vetores como entrada também usam uma rede neural convolucional 1D (CNN) para classificar as arritmias. Outros usam classificadores que combinam rede neural convolucional e rede de memória de curto prazo (CNN-LSTM), ou classificadores clássicos, como floresta aleatória, Naïve-Bayes ou rede neural rasa.

Para construir imagens binárias 2D, são usadas imagens de sinal de ECG. Conforme mostrado na Figura 1, essas imagens correspondem aos batimentos do ECG. Em Zheng *et al.*, os batimentos compreendem 92 amostras antes e depois do pico-R. A imagem binária 2D é usada como entrada dos classificadores CNN2D ou CNN2D-LSTM.

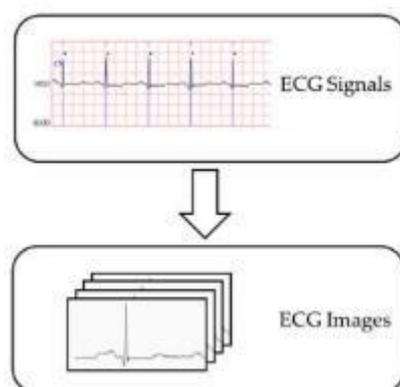


Figura 1. Extração de imagens direto do ECG.  
Fonte: Zheng *et al.*, 2020(adaptada).

Para a construção de espectrogramas 2D, os sinais unidimensionais de séries temporais de ECG são transformados em espectrogramas 2D por meio da transformada de Fourier de curta duração (Ullah *et al.*, 2020). Semelhante a este último trabalho, temos Ajjeey *et al.* (2020) que também utiliza espectrogramas 2D, neste trabalho os autores demonstram na Figura 2 o processo de transformação dos dados unidimensionais em espectrograma 2D através da transformada *wavelet* contínua (CWT). Nestes dois artigos, os autores usaram uma CNN 2D para classificação de arritmia.

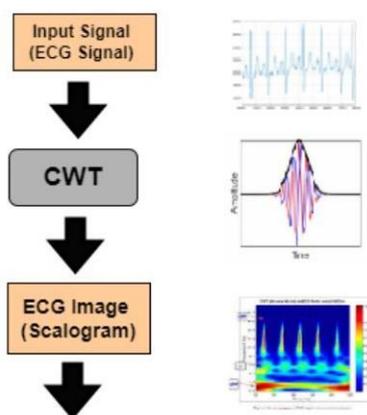


Figura 2. Transformação de sinais unidimensionais em espectrograma através da transformada wavelet contínua.  
 Fonte: AJJEY *et al.*, 2022 (adaptada)

Segundo Rohmantri *et al.*, (2020), a principal razão para usar sinais 2D em vez de sinais 1D para classificação de arritmia é que a CNN2D teve melhor desempenho na classificação de arritmia do que a CNN1D.

O banco de dados do MIT é muito desbalanceado. Por exemplo, as classes NSR, LBBB e RBBB possuem 75072, 8075 e 7259 anotações no banco de dados, respectivamente, enquanto as classes APB e PVC possuem apenas 150 e 82 anotações no banco de dados. Portanto, classificar as arritmias nas 17 classes é uma tarefa desafiadora. Os estudos apresentados na Tabela 1 classificam as arritmias em 2, 3, 4, 5, 7, 8, 10, 13, 15 ou 17 classes.

Buscando obter um conjunto de dados mais equilibrado, alguns trabalhos empregam diferentes técnicas de aumento de dados. Zheng *et al.*, (2020) empregaram 9 recortes de imagem diferentes em imagens digitais 2D de batimentos de ECG. Ullah *et al.*, (2020) também utilizaram corte de imagem em imagens de espectrogramas 2D. Qureshi *et al.*, (2022) empregaram a técnica de sobreamostragem minoritária sintética (SMOTE) em vetores de amostra com duração de 0,65 segundos. A técnica SMOTE é baseada no algoritmo de classificação *K-nearest neighbors* (KNN). Uma vez que dois

vizinhos estão conectados, um ponto aleatório é tomado ao longo desta conexão no espaço de características e um exemplo sintético é criado neste ponto (Abdalla *et al.*, 2022). Este procedimento é repetido até que a classe minoritária tenha tantas amostras em seu conjunto quanto a classe majoritária.

Na área médica, as decisões estão diretamente relacionadas à saúde dos pacientes. Portanto, é difícil para a equipe médica confiar nos modelos de aprendizado de máquina apenas com base em valores de alta precisão, sem saber a justificativa para a decisão. Para usar modelos de aprendizado de máquina para detecção de arritmia, é relevante informar à equipe médica sobre a base de decisão dos modelos e quais características dos dados inseridos nos modelos influenciam os resultados. Nos artigos mostrados na Tabela 1, apenas o estudo de Kim *et al.* utilizaram uma técnica para justificar a decisão do classificador, o mapeamento de ativação de classe ponderado por gradiente (KIM *et al.*, 2021), que foi desenvolvido para a interpretação visual de um modelo CNN. O Grad-CAM destaca no sinal de entrada, as regiões que têm maior influência na decisão do classificador. Usando uma rede DenseNet com uma arquitetura modificada, os autores mostraram que a classificação da arritmia é realizada pelo modelo usando uma parte específica do ECG. Na literatura, o estudo de Singh *et al.*, (2022) também aplicou técnicas para justificar a decisão de classificação. Os achados da classificação foram interpretados por meio dos métodos *SHapley Additive exPlanations* (SHAP), explicações localmente interpretáveis de modelos agnósticos (LIME), Grad-CAM e um novo método proposto pelos autores, oK-GradCam. Segundo os autores SHAP e LIME realizam perturbações aleatórias que às vezes produzem explicações não confiáveis. O método proposto compartilha 71% de similaridade com o SHAP e 81% de similaridade com os métodos Grad-CAM.

Em Singh *et al.*, (2022) o principal objetivo dos autores é avaliar o desempenho de técnicas para justificar a decisão na classificação de arritmias. Os autores também usaram o conjunto de dados MIT BIH. Eles classificaram as arritmias em seis classes: NSR, LBBB, RBBB, APB, PVC e PR. O classificador proposto foi um modelo CNN composto por camadas de convolução e um módulo residual para capturar as propriedades e correlações complexas dos dados de séries temporais de ECG. O sinal de entrada da CNN foi um complexo QRS, com 90 amostras antes do pico R e 170 amostras após o pico R. Os picos R foram detectados usando o algoritmo Pan Tompkins (PAN, 1985). O sinal de ECG foi normalizado entre 0 e 1. O conjunto de dados é composto por 10.000 batimentos de cada classe. Para utilizar técnicas para justificar a decisão de classificação em sinais

1D, os autores desenvolveram um método de mascaramento. As janelas do ECG foram avaliadas mascarando-as em 4 a 100 máscaras.

Diante das considerações levantadas, é verificado que na literatura atual não existem trabalhos que utilizam redes CNN2D com imagens 2D em escalas de cinza, geradas a partir da conversão dos valores extraídos do ECG, como visto os trabalhos ora usam segmentos de imagens extraídos direto do ECG ou usam imagens extraídas de espectrogramas. Também há poucos trabalhos que buscaram justificar a decisão da rede com a técnica GradCAM. Em relação ao uso de técnicas de aumento de dados, os poucos trabalhos existentes utilizam técnicas de aumento de dados convencionais, ao passo que em nosso trabalho propomos uma nova abordagem.

### **3. REFERENCIAL TEÓRICO**

Neste capítulo serão abordados os principais conceitos necessários ao entendimento deste trabalho. A seção 3.1 discorre sobre a anatomia básica do coração humano, assim como os fluxos sanguíneos. A seção 3.2 apresenta sobre o sinal cardíaco. A seção 3.3 apresenta sobre os eletrocardiogramas. A seção 3.4 discorre sobre as arritmias cardíacas. A seção 3.5 introduz os conceitos básicos de redes neurais artificiais. Por fim, na seção 3.6, aborda-se os conceitos de redes neurais convolucionais que é o foco deste trabalho.

#### **3.1 O Coração Humano**

O coração é um órgão muscular que bombeia o sangue através do sistema circulatório do corpo. Segundo o *National Heart, Lung, and Blood Institute* (2021), "o coração humano é do tamanho de cerca de um punho fechado e pesa cerca de 300 gramas". Ele está localizado no centro do peito, ligeiramente inclinado para a esquerda, e é cercado pelo saco pericárdico, que é um saco de tecido fibroso que ajuda a protegê-lo contra choques e danos externos. É composto por quatro câmaras: duas câmaras superiores, átrios ou aurículas e as duas câmaras inferiores são os ventrículos. As aurículas são responsáveis por receber o sangue de volta ao coração, enquanto os ventrículos são responsáveis por bombear o sangue para o resto do corpo. Sua anatomia pode ser observada conforme a Figura 3:

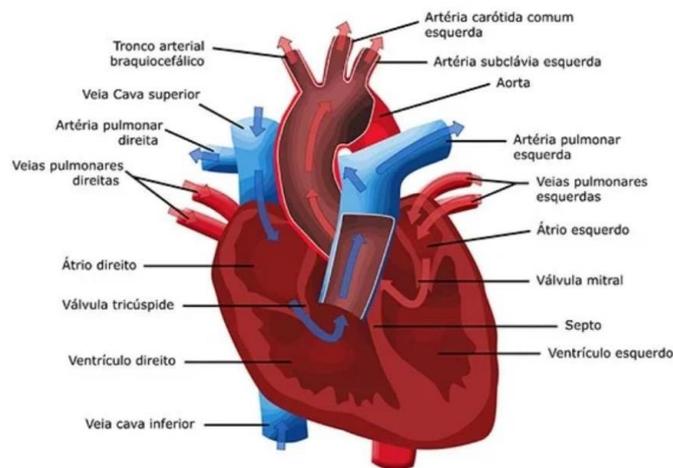


Figura 3. Anatomia do coração.

Fonte: Toda Matéria (2013).

Segundo Tortora e Derrickson (2017), A circulação pulmonar e a circulação sistêmica são processos fundamentais na distribuição de oxigênio e nutrientes no corpo humano. Na circulação pulmonar, o sangue é bombeado a partir do ventrículo direito para as artérias pulmonares. Essas artérias levam o sangue para os pulmões, onde ocorre a troca gasosa, permitindo que o sangue seja oxigenado. Em seguida, o sangue oxigenado retorna para o coração, através das veias pulmonares, indo para o átrio esquerdo. Na circulação sistêmica, o sangue oxigenado é bombeado a partir do átrio esquerdo para o ventrículo esquerdo, que o envia para todo o corpo através da artéria aorta. As células do corpo usam o oxigênio do sangue e produzem dióxido de carbono, que é então transportado pelo sangue de volta para o coração. O sangue retorna ao coração através das veias cavas superior e inferior e é levado ao átrio direito. Na Figura 4, é possível observar os sentidos dos fluxos sanguíneos dentro do coração:

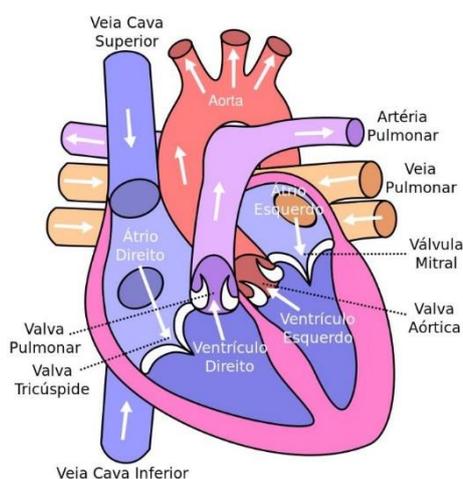


Figura 4. Fluxos sanguíneos dentro do coração.

Fonte: CEPED (2013).

### 3.2 Sinal Cardíaco

Os sinais originados das batidas do coração são iniciados com a contração dos átrios, em um movimento conhecido como sístole. Em seguida, as válvulas cardíacas se abrem e o sangue é bombeado para os ventrículos, que estão relaxados, em um movimento chamado de diástole. Esses movimentos compõem o ciclo cardíaco, gerando assim um sinal elétrico com um padrão específico (MAGALHÃES, 2018). Esse sinal elétrico se propaga pelo corpo, gerando atividade elétrica sob a pele e provocando tensão em determinadas partes do corpo. Quando eletrodos são colocados sobre a pele em pontos opostos do coração, os potenciais elétricos gerados por esses sinais podem ser registrados, resultando em um eletrocardiograma (ECG).

Um sinal de ECG é composto por diversos componentes, como ilustrado na Figura 5. A onda P corresponde à despolarização atrial e possui uma primeira componente relativa à aurícula direita e uma segunda componente relativa à aurícula esquerda. O complexo QRS representa a despolarização ventricular, sendo maior que a onda P, devido à massa muscular dos ventrículos ser maior do que a dos átrios, resultando em sinais de despolarização ventricular mais intensos do que os da repolarização atrial. Por fim, temos a onda T, que corresponde à polarização atrial. Os trechos entre cada segmento são representados pelo segmento PR, que vai da onda P ao pico R, o segmento ST, que vai do vale S até a onda T, e o segmento QT, que vai do vale Q até a onda T.

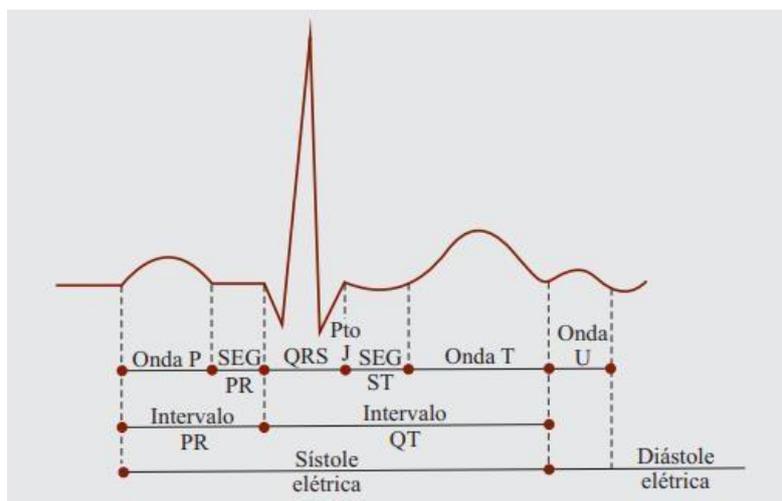


Figura 5. Caracterização esquemática do ECG com todas as suas ondas (P, Q, R, S, T e U) e seus intervalos (PR, QT e ST).  
Fonte: Atheneu (2013).

### 3.3 Eletrocardiograma

O eletrocardiógrafo é um dispositivo que amplifica, filtra e registra a atividade elétrica do coração em um papel milimetrado projetado especificamente para essa finalidade. De forma mais precisa, o registro consiste na diferença de potencial elétrico captada por eletrodos posicionados sobre a superfície do corpo de um indivíduo. Essa técnica permite a visualização e análise das características elétricas do coração, auxiliando no diagnóstico e monitoramento de condições cardíacas (Reis *et al.*, 2013). A Figura 6 apresenta um exemplo de um eletrocardiógrafo.



Figura 6. Eletrocardiógrafo.  
Fonte: Alfamed (2022).

De acordo com o livro "ECG: manual prático de eletrocardiograma" (Reis *et al.*, 2013), o traçado do ECG (eletrocardiograma) é representado por ondas que possuem características individuais, como duração, amplitude e configuração. A velocidade-padrão do papel milimetrado é de 25 mm/s, onde um quadrado menor corresponde a 0,04 s e um quadrado maior corresponde a 0,2 s, permitindo a determinação da duração dos eventos registrados. Em relação à amplitude dos traçados eletrocardiográficos, cada linha vertical é equivalente a 0,1 mV, demonstrado na Figura 7, a seguir.

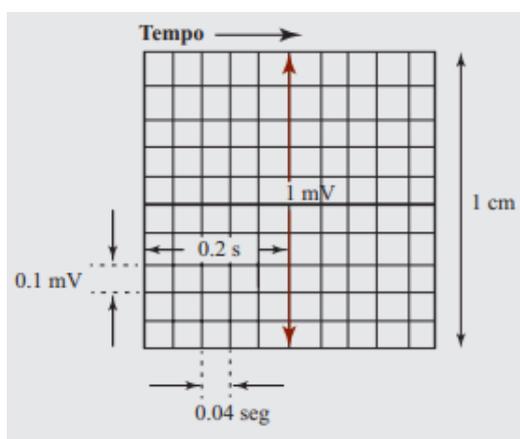


Figura 7. Representação do papel milimetrado do ECG com a duração em segundos na linha horizontal e a amplitude em milivolts na linha vertical.  
Fonte: Atheneu (2013).

De acordo com Silva *et al.* (2015), existem três abordagens para a medição de um sinal cardíaco: *in-the-person*, *on-the-person* e *off-the-person*. Na abordagem *in-the-person*, são utilizados equipamentos implantados cirurgicamente ou ingeridos por meio de pílulas. Na abordagem *on-the-person*, os aparelhos são conectados no corpo do paciente, como os eletrodos, permitindo a leitura dos sinais elétricos na pele. Já na abordagem *off-the-person*, há um mínimo contato com o corpo humano, e exemplos de equipamentos incluem dispositivos baseados em capacitores que medem os sinais elétricos a cerca de 1 cm da pele. A maioria dos aparelhos utiliza a abordagem *on-the-person* para realizar a medição, esta abordagem é mostrada na Figura 8.

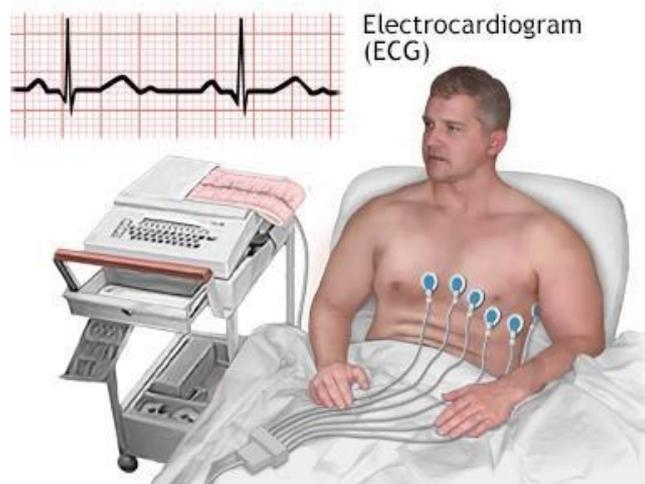


Figura 8. Leitura de sinais elétricos no método *on the person*.  
Fonte: SEDIC (2023).

Quanto à utilização de eletrodos para a abordagem *on-the-person*, existem diferentes configurações para posicionar esses dispositivos no corpo do paciente com o objetivo de realizar a medição do ritmo cardíaco. Segundo Bahia (2019), são empregadas cinco localizações específicas para o posicionamento dos eletrodos, nomeadamente: Braço Esquerdo (LA), Braço Direito (RA), Perna Esquerda (LL), Perna Direita (RL) e a direita do Esterno (V ou V1). Além disso, é possível encontrar a utilização de um conjunto de dez eletrodos para essa medição, no qual, além das cinco posições mencionadas anteriormente, outros cinco eletrodos são colocados no tórax do paciente (V2 a V6). Essas configurações podem ser visualizadas na Figura 9.

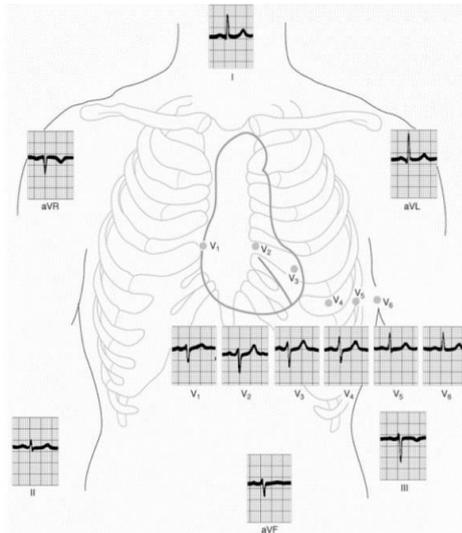


Figura 9. Pontos de coleta de frequências cardíacas em um ECG.  
Fonte: BAHIA (2019).

### 3.4 Arritmias Cardíacas

Segundo a Sociedade Brasileira de Arritmias Cardíacas (SOBRAC), as arritmias cardíacas são alterações que afetam a geração ou condução do estímulo elétrico do coração, resultando em modificações no ritmo dos batimentos cardíacos. Essas condições podem ocorrer em pessoas de todas as faixas etárias, desde crianças até idosos, e são responsáveis por muitas mortes súbitas no Brasil, com uma média de cerca de trezentas mil por ano. As arritmias podem ser causadas por diversos fatores, tais como desaceleração ou aceleração anormal do ritmo cardíaco, bloqueio na passagem do sinal elétrico no nodo atrioventricular ou feixe atrioventricular, presença de sinais reentrantes em áreas isquêmicas, efeitos de substâncias como cafeína e nicotina, além de calcificações no tecido cardíaco, entre outros.

Segundo a *American Heart Association* (2017), as arritmias cardíacas podem levar a uma redução na eficiência da bomba cardíaca. Elas podem ser classificadas em dois tipos principais: taquiarritmias, que são arritmias com aumento da frequência cardíaca, e bradiarritmias, que são arritmias com diminuição da frequência cardíaca. Alguns sintomas das arritmias cardíacas incluem palpitações, tonturas, desmaios e falta de ar.

A seguir a Figura 10 exibe um segmento de um ECG que ilustra um exemplo de arritmia conhecida como batida de bloqueio de ramo direito (R), enquanto a Figura 11 exibe um ECG contendo a arritmia conhecida como contração ventricular prematura (V).

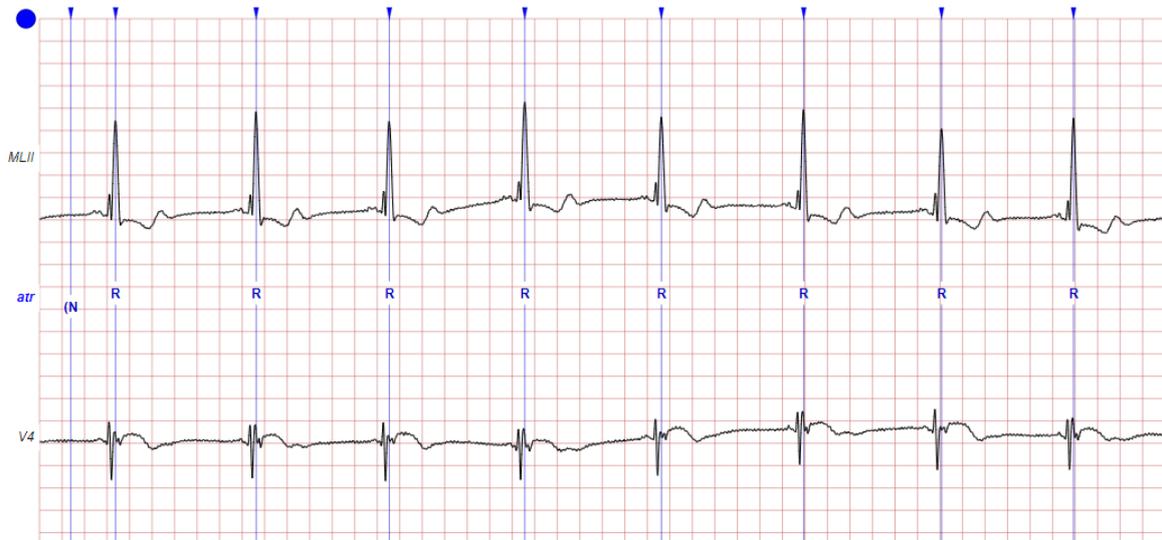


Figura 10. Trecho de um exame de ECG contendo a arritmia batida de bloqueio de ramo direito (R).  
 Fonte: PHYSIONET (2018).

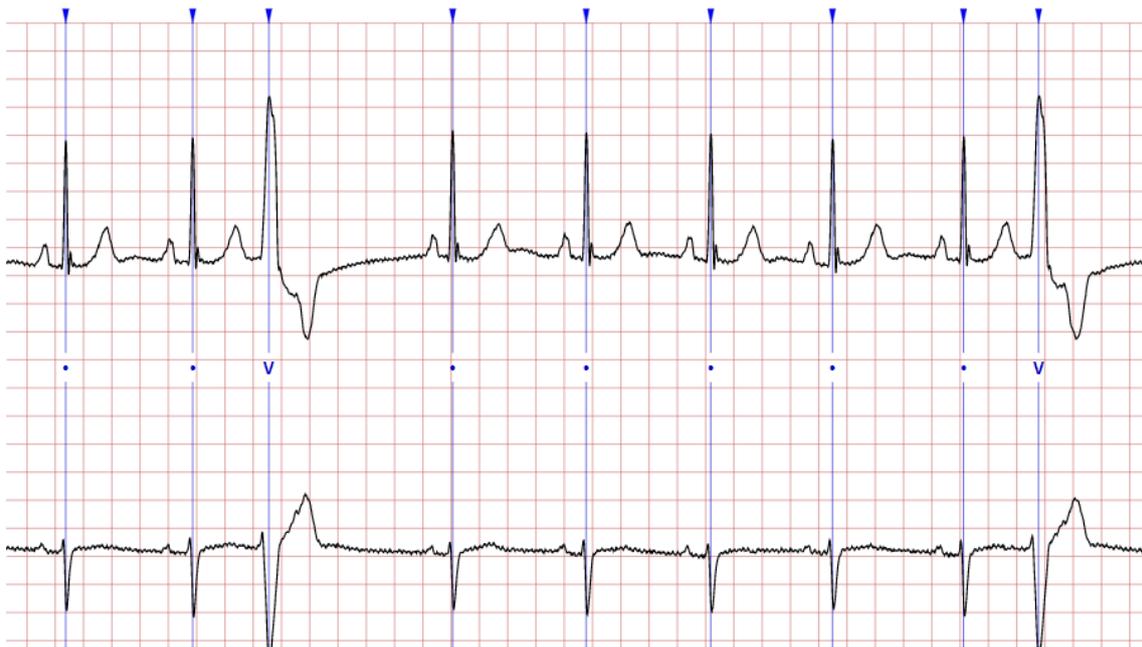


Figura 11. Trecho de um exame de ECG contendo a arritmia contração ventricular prematura (V).  
 Fonte: PHYSIONET (2018).

As 13 classes, cujos sinais encontram-se disponíveis na base de dados do MIT-BIH (PHYSIONET, 2018) utilizadas neste trabalho são:

- 1) Batimento de bloqueio de ramo esquerdo (*Left bundle branch block beat – L*);
- 2) Batida de bloqueio de ramo direito (*Right bundle branch block beat – R*);
- 3) Batimento atrial prematuro aberrado (*Aberrated atrial premature beat – a*);
- 4) Batimento prematuro nodal (*Nodal premature beat – J*);
- 5) Contração ventricular prematura (*Premature ventricular contraction – V*);
- 6) Fusão de batimento ventricular e normal (*Fusion of ventricular and normal beat – F*);
- 7) Batida de fuga atrial (*Atrial escape beat – e*);
- 8) Batida de fuga nodal (*Nodal escape beat – j*);
- 9) Batida de escape ventricular (*Ventricular escape beat – E*);
- 10) Fusão de batimento compassado e normal (*Fusion of paced and normal beat – f*);
- 11) Batida inclassificável (*Unclassifiable beat – Q*);
- 12) Ritmo compassado (*paced rhythm – P*);
- 13) Batimento normal (*Normal beat – N*).

### 3.5 Redes Neurais Artificiais (RNA)

Para compreender o funcionamento das redes neurais, é essencial ter conhecimento sobre a dinâmica do cérebro humano. Segundo Costa e Ludwig Júnior (2007), os neurônios são as unidades fundamentais do cérebro humano, responsáveis pela transmissão de informações e condução de mensagens nervosas. Um neurônio é composto por três partes principais: o soma ou corpo celular, dendritos e axônio. Nos terminais dos axônios, ocorrem as transmissões de informações para outros neurônios, por meio das sinapses. A Figura 12 ilustra a estrutura de uma célula neuronal.

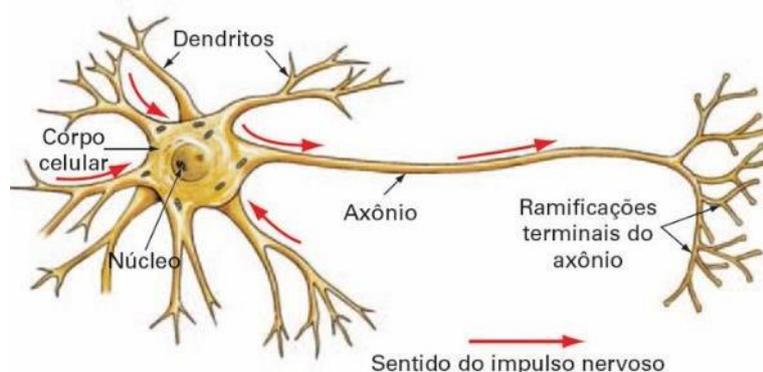


Figura 12. Representação Simplificada do Neurônio Biológico.

Fonte: KHAN ACADEMY (2021).

No corpo da célula neuronal, ocorre o processamento das informações recebidas de outras células nervosas, e essas informações são enviadas para outros neurônios. Os dendritos desempenham o papel de receber os impulsos nervosos provenientes de outros neurônios e conduzi-los até o corpo celular. Dessa forma, a informação coletada pelos dendritos é enviada e processada no corpo do neurônio. O corpo celular realiza a soma dessas informações e gera pulsos elétricos com base no valor do potencial elétrico resultante desse processo de soma.

O pulso ou potencial de ação só é gerado pelo neurônio se o potencial elétrico resultante da soma das informações for superior a um limiar de  $-50\text{mV}$ , conforme mencionado por Costa e Ludwig Júnior (2007). Esses pulsos gerados são transmitidos através do axônio para os dendritos dos neurônios subsequentes.

No campo da Inteligência Artificial, o modelo de um neurônio artificial, inspirado no funcionamento do neurônio biológico, tornou-se a base para o desenvolvimento de sistemas inteligentes. Segundo Costa e Ludwig Júnior (2007), o modelo mais amplamente aceito foi proposto por Warren McCulloch e Walter Pitts em 1943, e apresentava uma representação simplificada dos componentes e do funcionamento de um neurônio biológico.

Nesse modelo, o neurônio artificial realiza o cálculo da soma ponderada das diversas entradas e submete o resultado a uma função de ativação, também conhecida como função de transferência. A saída do neurônio é determinada pela função de ativação aplicada ao resultado da soma ponderada. A Figura 13 ilustra a arquitetura básica de um neurônio artificial.

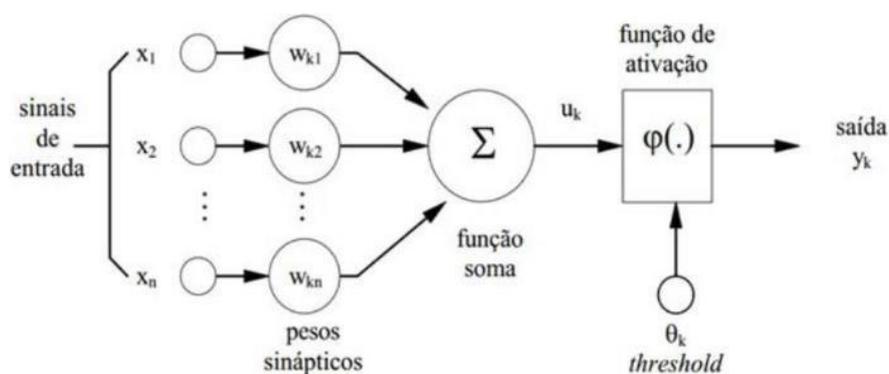


Figura13. Modelo de um Neurônio Artificial.  
Fonte: HAYKIN (2001).

Assim como no neurônio biológico, o neurônio artificial pode receber estímulos

excitatórios ou inibitórios em suas entradas. No modelo de McCulloch e Pitts, esses estímulos são representados por valores numéricos conhecidos como pesos sinápticos. Valores positivos dos pesos sinápticos representam estímulos excitatórios, enquanto valores negativos indicam estímulos inibitórios. A magnitude de um peso sináptico determina sua importância e relevância na computação realizada pelo neurônio.

No neurônio artificial a relação entre a entrada e a saída é mostrada através da equação (1) e no formato matricial, na (2):

$$X.W = y \quad (1)$$

em que:

X = Vetor das entradas;

W = Matriz dos pesos e polarizações;

y = Vetor das saídas.

$$\begin{bmatrix} 1 & X_{11} & \dots & X_{1d} \\ 1 & X_{21} & \dots & X_{2d} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & X_{n1} & \dots & X_{nd} \end{bmatrix} \times \begin{bmatrix} W_0 \\ W_0 \\ \vdots \\ W_d \end{bmatrix} = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_d \end{bmatrix} \quad (2)$$

É importante ressaltar que a representação dos neurônios artificiais e das redes neurais em sistemas de Inteligência Artificial permite a realização de cálculos complexos e a capacidade de aprendizado a partir dos dados de entrada. As redes neurais artificiais são compostas por múltiplos neurônios interconectados, que trabalham em conjunto para processar informações e fornecer respostas ou previsões baseadas em padrões identificados nos dados.

O uso de funções de ativação não-lineares é fundamental para potencializar a capacidade das redes neurais artificiais. Essas funções introduzem não-linearidades nos cálculos realizados pelos neurônios, permitindo a modelagem de relações complexas entre as entradas e as saídas da rede. Essa propriedade não-linear das funções de ativação torna as redes neurais poderosas em termos de sua capacidade de aprendizado e representação de dados.

Além das funções de ativação, os modelos de neurônios artificiais incluem uma variável chamada polarização ou *bias*. Essa variável é adicionada ao somatório da função de ativação e tem o objetivo de aumentar o grau de liberdade da função e melhorar a capacidade de aproximação da rede. Assim como os pesos sinápticos, o valor do *bias* também é ajustado durante o processo de treinamento da rede. A inclusão do *bias*

possibilita que um neurônio tenha uma saída não nula, mesmo quando todas as suas entradas são nulas. Isso é importante para garantir que a função de ativação não seja anulada quando todas as entradas do neurônio são zero (COSTA e LUDWIG JÚNIOR, 2007).

### 3.5.1 Funções de Ativações

Segundo Bishop (1995), a função de ativação desempenha um papel crucial nas redes neurais, pois introduz não-linearidades que permitem à rede aprender representações complexas e generalizadas dos dados. Além disso, a função de ativação deve ser diferenciável para permitir o treinamento eficiente do modelo. São mostradas nas equações (3), (4), (5) e (6):

$$f(x) = \frac{e^{+x} - e^{-x}}{e^{+x} + e^{-x}} \quad (3)$$

$$f(x) = \frac{1}{1 + e^{-x}} \quad (4)$$

$$f(x) = \max(0, x) \quad (5)$$

$$S_j = \frac{z_j}{\sum_{k=1}^k e^{z_k}} \quad (6)$$

em que:

k = índice dos neurônios do mesmo nível;

z = vetor dos neurônios de saída;

j = índice do neurônio de saída.

É importante ressaltar que cada função de ativação tem suas propriedades e é mais adequada para determinados tipos de problemas. A escolha da função de ativação depende da natureza do problema e das características dos dados.

### 3.5.2 Perceptron Multicamadas

Devido à simplicidade do modelo do *perceptron*, algumas limitações foram identificadas, sendo a principal delas a sua incapacidade de classificar classes que não

sejam linearmente separáveis (MINSKEY E PAPERT, 1969). Para superar essa limitação, uma abordagem mais avançada é o uso de redes neurais de múltiplas camadas, conhecidas como MLP (do inglês *Multilayer Perceptron*). Essa arquitetura consiste em uma ou mais camadas ocultas, compostas por vários neurônios. A inclusão das camadas ocultas permite que o MLP aprenda representações mais complexas dos dados, tornando-se capaz de lidar com problemas de classificação não linear. Para uma melhor compreensão dessa rede, a Figura 14 ilustra um exemplo de um MLP com duas camadas ocultas.

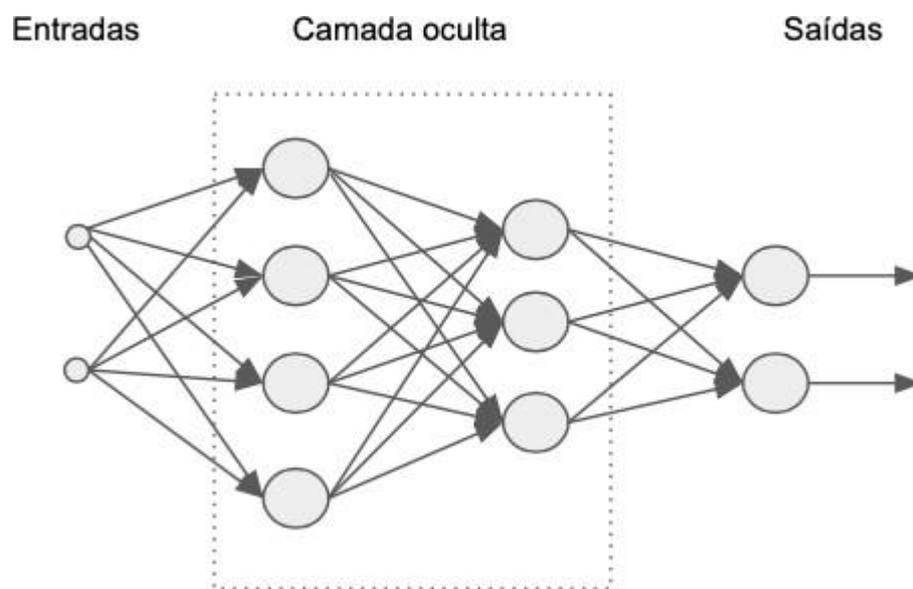


Figura 14. Ilustração de um MLP com duas camadas ocultas.

Fonte: ATELIWARE (2021).

### 3.5.3 Algoritmo Backpropagation

O algoritmo de *Backpropagation*, também conhecido como retropropagação do erro, é um dos principais algoritmos utilizados no treinamento de redes neurais artificiais, especificamente em redes do tipo *Multilayer Perceptron* (MLP). Ele é responsável por ajustar os pesos sinápticos da rede de forma a minimizar o erro entre a saída desejada e a saída produzida pela rede.

O algoritmo de *Backpropagation* é composto por duas etapas: a propagação direta (do inglês *forward propagation*) e a propagação reversa (do inglês *backward propagation*). Na etapa de propagação direta, os dados de entrada são alimentados na rede, e os sinais são propagados camada por camada até a camada de saída, onde é gerada a saída predita pela rede. Durante essa etapa, os valores de ativação de cada neurônio são

calculados com base nos pesos sinápticos atuais.

Na etapa de propagação reversa, o erro entre a saída predita e a saída desejada é calculado. Esse erro é então propagado de volta pela rede, camada por camada, para ajustar os pesos sinápticos. Esse ajuste é feito utilizando o algoritmo do gradiente descendente, que calcula a taxa de variação do erro em relação aos pesos e atualiza os pesos na direção oposta ao gradiente, buscando minimizar o erro.

O processo de propagação reversa é repetido iterativamente para cada exemplo de treinamento, ajustando gradualmente os pesos da rede até que o erro seja minimizado e a rede seja capaz de produzir saídas mais precisas.

O algoritmo de *Backpropagation* é fundamental para o treinamento de redes neurais MLP e tem sido amplamente utilizado em diversas aplicações, como reconhecimento de padrões, processamento de sinais, previsão e muitas outras áreas. (HAYKIN, 2008).

### **3.6 Redes Neurais Convolucionais (CNN)**

Entre os modelos de redes neurais artificiais, as Redes Neurais Convolucionais (CNNs) são amplamente utilizadas em tarefas de processamento de imagens e reconhecimento visual (JONES, 2017).

No contexto de tarefas de classificação, as CNNs têm sido especialmente benéficas. Enquanto abordagens tradicionais requeriam a extração manual de características dos sinais a serem classificados, as CNNs são capazes de aprender essas características por conta própria. Isso elimina a necessidade de uma etapa de pré-processamento e extração de características, tornando o processo mais eficiente e automático. As CNNs são capazes de identificar padrões complexos e sutis nos dados de entrada, permitindo uma classificação mais precisa e robusta.

A arquitetura de uma Rede Neural Convolucional (CNN) é composta por diversas camadas, cada uma desempenhando uma função específica no processamento dos dados de entrada. A Figura 15 ilustra uma possível estrutura de uma rede CNN.

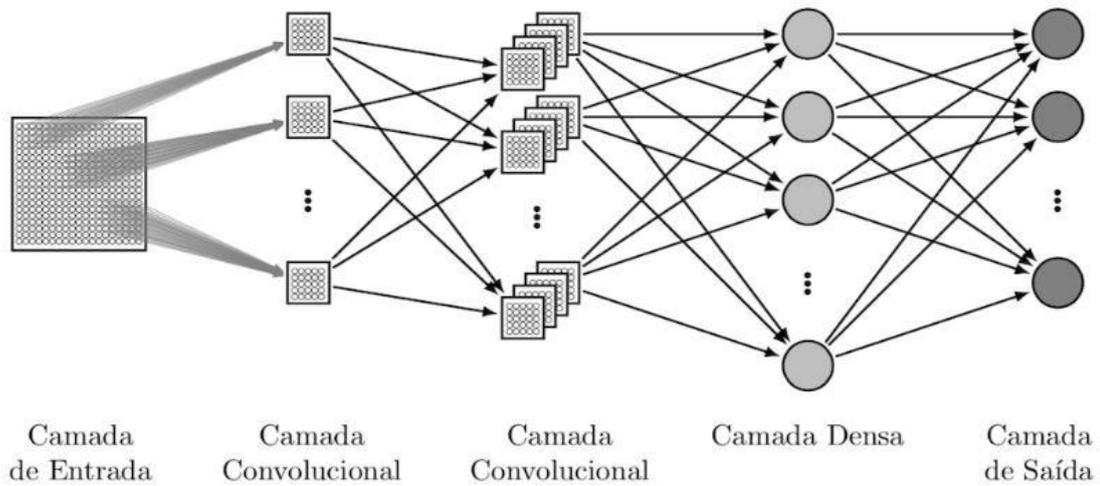


Figura 15. Modelo descritivo de uma rede neural convolutiva.  
 Fonte: SAKURAI (2017).

### 3.6.1 Convoluções

Conforme documentado na literatura, uma das principais camadas das redes CNNs é a camada de convolução. Nas redes CNNs a Camada Convolucionacional ou de Convolução utiliza filtros para obter as características mais expressivas dos dados de entradas da CNN. Os filtros são compostos por janelas deslizantes que se movimentam através do campo receptivo para a extração de características. A característica a ser identificada é definida através do tipo de filtro (*kernel*) a ser utilizado, como por exemplo, o filtro de identificação de bordas. Este filtro percorre a imagem toda, realizando operações de convolução sobre a região chamada de campo receptivo e guardando os resultados obtidos no mapa de ativações ou, como também é chamado, mapa de parâmetros (DESHPANDE, 2017). Dessa forma, gera-se um mapa de características. Ela é definida matematicamente pela equação (7). Nessa equação,  $I(x, y)$  representa a imagem,  $f(x, y)$  é o filtro ou *kernel*, o qual tem  $a$  e  $b$  como suas dimensões.

$$g(x, y) = I(x, y) * f(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b I(s, t) * f(x - s, y - t) \quad (7)$$

Na Figura 16, é mostrada uma operação de convolução.

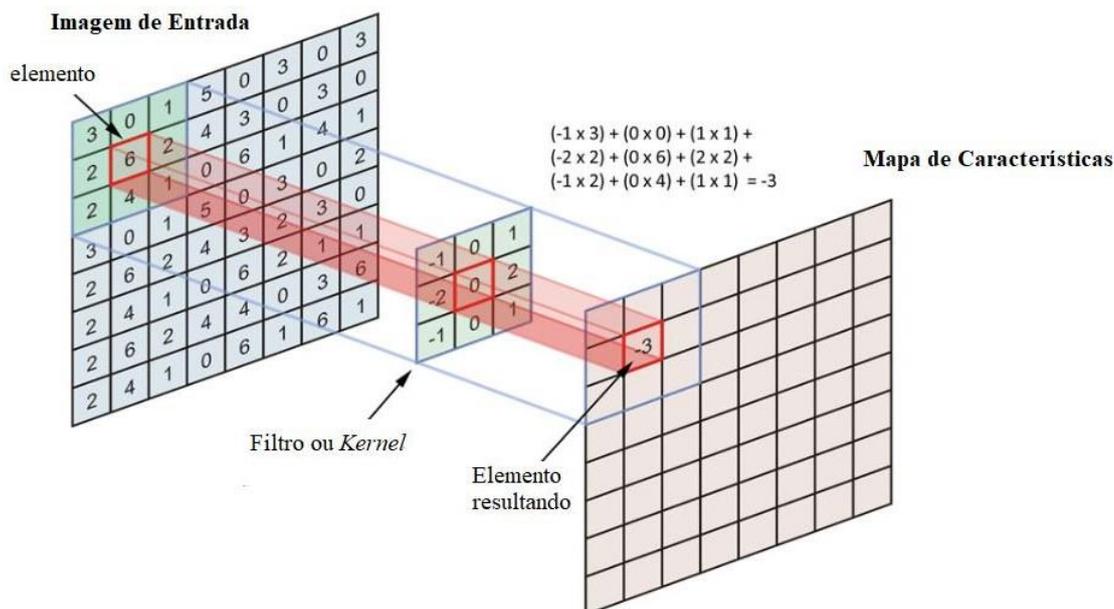


Figura 16. Exemplo de operação de convolução.  
Fonte: LAMFO (2020).

Nas redes CNNs, é comum usar uma camada de amostragem, também conhecida como camada de *pooling*, para simplificar as informações das camadas anteriores.

A camada de *pooling* é estruturada de forma similar à camada de convolução. Ela opera em uma pequena região de neurônios da camada anterior, chamada de campo receptivo, e produz um único valor real com base nessa região.

A operação mais comumente utilizada na camada de *pooling* é o *maxpooling*, em que o valor máximo da região é selecionado como representante daquela região. No entanto, também existem outras variações, como o *average pooling*, em que o valor médio é calculado.

A principal função da camada de *pooling* é reduzir as dimensões espaciais (largura e altura) da representação da imagem. Isso é feito pela redução da resolução, mantendo as características mais importantes. Ao reduzir as dimensões, a quantidade de parâmetros e computações necessárias nas camadas subsequentes é diminuída, o que ajuda a evitar o *overfitting* e a reduzir o custo computacional.

Além da redução de dimensões, a camada de *pooling* também introduz uma certa invariância a translações e rotações. Isso significa que mesmo que um objeto esteja ligeiramente deslocado ou rotacionado na imagem, a camada de *pooling* ainda será capaz de identificar as características relevantes, desde que elas estejam presentes na região de campo receptivo. Na Figura 17 mostramos a operação de *maxpooling*.

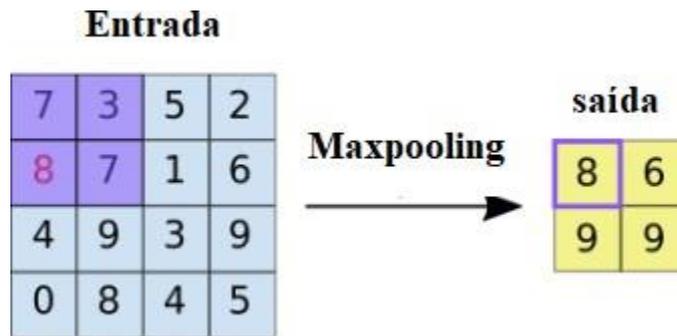


Figura 17. Operação de *maxpooling*.  
Fonte: PEIXOTO (2020).

Nas CNNs, as camadas totalmente conectadas (do inglês *Fully Connected* -FC) são comumente usadas como camadas finais da rede, mas não necessariamente. Essas camadas são estruturas lineares de neurônios, onde cada neurônio está conectado a todos os neurônios da camada anterior.

Se a camada FC for a última camada da rede, o número de neurônios nessa camada geralmente corresponderá à quantidade de classes do problema de reconhecimento de padrões que está sendo resolvido. Cada neurônio na camada de saída representará uma classe específica e a ativação do neurônio indicará a probabilidade de pertencer a essa classe.

### 3.6.2 Camada de *Dropout*

Com o objetivo de reduzir os efeitos do *overfitting* em redes neurais convolucionais (CNNs), algumas técnicas podem ser aplicadas, como o *dropout*, a regularização L1 e L2 e a parada antecipada. O *overfitting* ocorre quando a rede se adapta excessivamente aos dados de treinamento, levando a uma baixa capacidade de generalização para novos dados (GOODFELLOW, BENGIO e COURVILLE, 2016).

O *dropout* é uma técnica amplamente utilizada para reduzir o *overfitting* em redes neurais, incluindo CNN. Essa técnica consiste em desativar aleatoriamente algumas unidades (neurônios) da rede durante o treinamento, fazendo com que essas unidades não contribuam para o cálculo dos gradientes e para a propagação do erro. Isso ajuda a evitar que unidades específicas se tornem muito dependentes umas das outras e reduz a possibilidade de *overfitting*, pois a rede se torna mais robusta e generalizável.

Ao utilizar a técnica de *dropout* em arquiteturas de CNN, algumas unidades da rede são desativadas aleatoriamente durante o treinamento, como ilustrado na Figura 18. Isso reduz a capacidade da rede de se ajustar excessivamente aos dados de treinamento,

tornando-a mais robusta e melhorando sua capacidade de generalização.

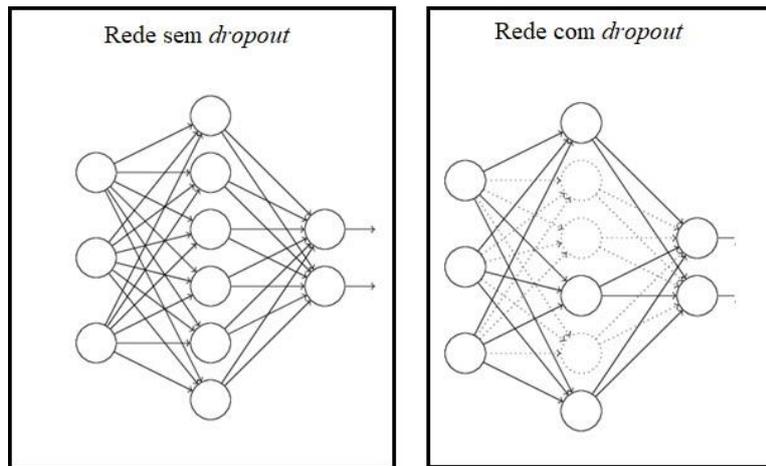


Figura 18. Resultado da aplicação de *dropout*.  
Fonte: Yongjie *et al.* (2018).

#### 4. MATERIAIS E MÉTODOS

A metodologia proposta neste trabalho para a identificação e classificação de arritmias cardíacas em sinais de eletrocardiograma utiliza diversas técnicas de processamento digital de imagens e de reconhecimento de padrões, dentre as quais destaca-se o reescalonamento da imagem, o aumento de dados e as redes neurais convolucionais 2D. A Figura 19 mostra o fluxograma da metodologia utilizada. É mister frisar que a ideia central deste trabalho é avaliar o desempenho de uma arquitetura CNN2D na classificação das arritmias cardíacas.

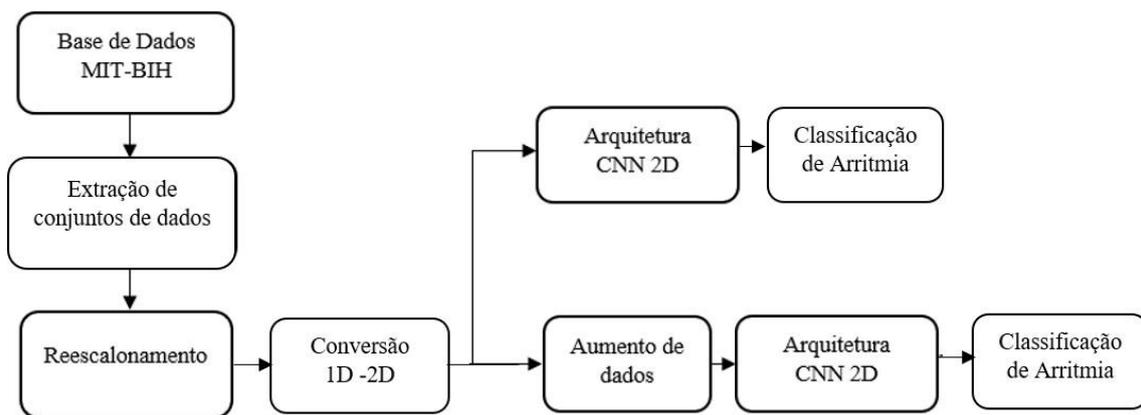


Figura 19. Fluxograma da metodologia utilizada para a identificação e classificação de arritmias cardíacas.

Conforme mostrado na Figura 19, o desempenho da arquitetura CNN2D é avaliado em duas etapas. Primeiramente, utilizando-se as amostras originais de ECG

contidas no banco de dados MIT- BIH (vide item 4.1). Posteriormente, utilizando amostras de um conjunto maior, obtido através de técnicas de aumento de dados. A ideia central do aumento de dados é possibilitar um melhor ajuste dos parâmetros da rede CNN2D, evitando o *overfitting* durante o treinamento dela.

#### 4.1 MATERIAIS

A base de dados escolhida para a extração dos sinais de ECG, foi a “MIT-BIH *Arrhythmia Database*” coletada pelo Laboratório de Arritmias do Hospital Beth Israel, em parceria com o Instituto de Tecnologia de Massachusetts. Essa base está disponível no banco de dados da *PhysioNet* (PHYSIONET, 2021).

Os sinais de ECG, obtidos a partir de 47 pacientes, foram amostrados a uma taxa de 360 Hz. No total, a base de dados é composta por 48 registros anotados, com duração aproximada de 30 minutos cada um. Na caracterização do banco, estes registros foram selecionados de forma aleatória e foram separados em dois intervalos. O primeiro intervalo é composto por 23 registros, iniciando no registro 100 e finalizando no registro 124 (não contém o registro 110). O segundo intervalo é composto por 25 registros, iniciando no registro 200 e finalizando no registro 234 (não contém os registros: 204, 206, 211, 216, 218, 224, 225, 226, 227 e 229) (PHYSIONET, 2021).

Define-se evento como a ocorrência de um batimento cardíaco. Os batimentos cardíacos podem ser classificados como batimentos normais ou batimentos anormais, também denominados de arritmias cardíacas. O número de eventos em um registro corresponde, então, ao número de batimentos cardíacos encontradas no mesmo. É possível, em um único registro, encontrar-se mais de um tipo de arritmia cardíaca. Na Tabela 2, mostramos alguns registros com suas respectivas quantidades de eventos anormais, nesse caso batimento prematuro atrial e contração ventricular prematura. Observa-se que em todos os registros mostrados na Tabela 2 existe pelo menos um tipo de arritmia.

Tabela 2. Exemplos de Registros da base MIT-BIH com número de eventos contabilizados. Observa-se que em cada registro existe pelo menos um tipo de arritmia.

Registro	Número de Eventos		
	Batida normal (N)	Batimento prematuro atrial (A)	Contração ventricular prematura (V)
100	2239	33	1
103	2082	2	--
116	2302	1	109
119	1543	--	444
221	2031	--	396
228	1688	3	362

Ao longo dos 48 registros da base MIT-BIH são identificadas 17 classes de eventos. Esses eventos, conforme já explicado, podem ser batimentos normais ou arritmias cardíacas (batimentos anormais). Na Tabela 3 apresenta-se essas classes e o número de eventos encontrados em cada uma. Ao analisar a coluna “número de eventos” da Tabela 3, é possível notar que existe um grande desbalanceamento entre classes. A classe de batida normal (N), por exemplo, possui 75.052 eventos. Em contrapartida, a classe batidade fuga atrial (e) apresenta apenas 16 ocorrências.

Tabela 3. As 17 classes de arritmia cardíaca contidas na base MIT-BIH com número de eventos contabilizados.

	Símbolo	Tipo de Evento	Número de Eventos
1	<b>N</b>	Batida normal	75.052
2	<b>L</b>	Batimento de bloqueio de ramo esquerdo	8.075
3	<b>R</b>	Batida de bloqueio de ramo direito	7.259
4	<b>a</b>	Batimento atrial prematuro aberrado	150
5	<b>J</b>	Batimento prematuro nodal (juncional)	83
6	<b>V</b>	Contração ventricular prematura	7.130
7	<b>F</b>	Fusão de batimento ventricular e normal	803
8	<b>e</b>	Batida de fuga atrial	16
9	<b>j</b>	Batida de fuga nodal (juncional)	229
10	<b>E</b>	Batida de escape ventricular	106
11	<b>f</b>	Fusão de batimento compassado e normal	982
12	<b>Q</b>	Batida inclassificável	33
13	<b>O</b>	Onda de vibração ventricular	665
14	<b>A</b>	Batimento prematuro atrial	2.546
15	<b>P</b>	Ritmo compassado	7.028
16	<b>/</b>	Batida compassada	45
17	<b>S</b>	Batimento supraventricular prematuro	2

A maioria dos trabalhos publicados na literatura que utilizam a base MIT-BIH destinam-se a classificação de 5, 8 ou 13 classes de arritmias (OH *et al.*, 2018, ZHENG *et al.*, 2020, YLDIRIM *et al.*, 2018). À semelhança do trabalho de YLDIRIM e colaboradores (2018), nesta dissertação faremos a classificação de 13 classes de arritmias, as quais são apresentadas na Tabela 4. Pode-se verificar ainda nessa Tabela que o número de eventos de algumas classes foi reduzido, podemos citar como exemplo a classe de Batimento de bloqueio de ramo esquerdo (L) que possui um número de eventos em 8.075 e foi reduzida para 1.000 visando minimizar problemas durante o treinamento, como por exemplo, a rede se especializar nessa classe por possuir um grande número de amostras. De semelhante forma esse critério foi aplicado as classes: Batida Normal (N), Batida de bloqueio de ramo direito (R), Contração ventricular prematura (V) e Ritmo compassado (P).

Tabela 4. Classes de arritmias cardíacas da base de dados MIT-BIH a serem utilizadas neste trabalho com a devida redução do número de eventos de algumas classes.

	<b>Símbolo</b>	<b>Tipo de Evento</b>	<b>Número de Eventos</b>
1	<b>N</b>	Batida normal	1.000
2	<b>L</b>	Batimento de bloqueio de ramo esquerdo	1.000
3	<b>R</b>	Batida de bloqueio de ramo direito	1.000
4	<b>a</b>	Batimento atrial prematuro aberrado	150
5	<b>J</b>	Batimento prematuro nodal (juncional)	83
6	<b>V</b>	Contração ventricular prematura	1.000
7	<b>F</b>	Fusão de batimento ventricular e normal	803
8	<b>e</b>	Batida de fuga atrial	16
9	<b>j</b>	Batida de fuga nodal (juncional)	229
10	<b>E</b>	Batida de escape ventricular	106
11	<b>f</b>	Fusão de batimento compassado e normal	982
12	<b>Q</b>	Batida inclassificável	33
13	<b>P</b>	Ritmo compassado	1.000
<b>Total</b>			<b>7.402</b>

O conjunto de dados utilizado neste trabalho engloba dados das 13 classes mostradas na Tabela 4 e foram obtidos através de duas etapas. A primeira etapa consistiu na extração de dados para a formação dos conjuntos de eventos, correspondentes a batimentos normais ou arritmias. A segunda etapa consistiu na formação de subconjuntos de eventos a partir dos conjuntos de cada classe. Cada subconjunto é constituído por 1600 amostras do ECG. Este processo é demonstrado através do fluxograma da Figura 20.

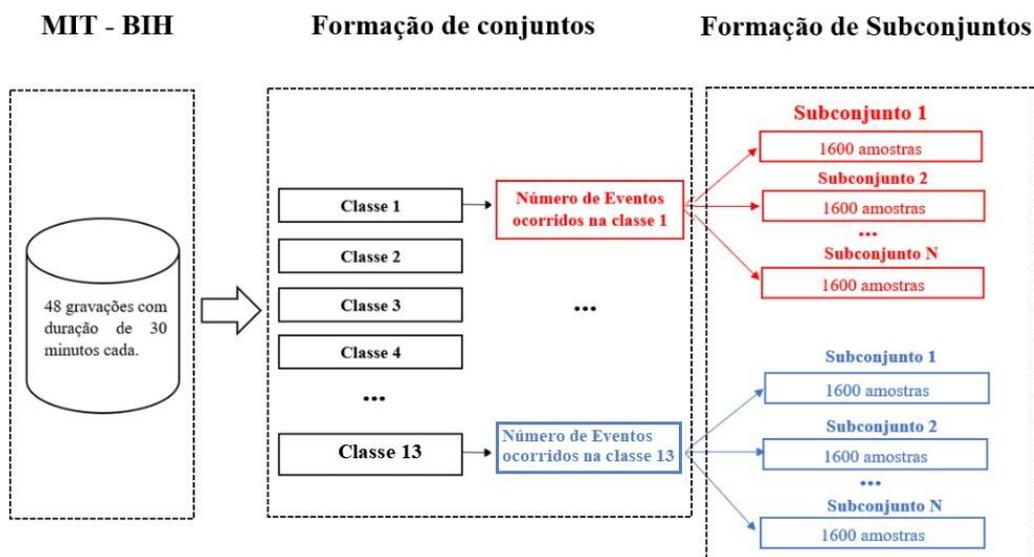


Figura 20. Diagrama em blocos mostrando o processo de extração e formação dos conjuntos de dados de cada classe e a formação de subconjuntos de 1600 amostras.

Nas Figuras 21, 22 e 23 são apresentados subconjuntos pertencentes às classes “a” (batimento prematuro atrial), “L” (bloqueio do ramo esquerdo) e “j” (batida de fuga nodal), respectivamente. O eixo horizontal é composto por 1600 amostras, enquanto o eixo vertical é composto por valores de amplitudes do ECG.

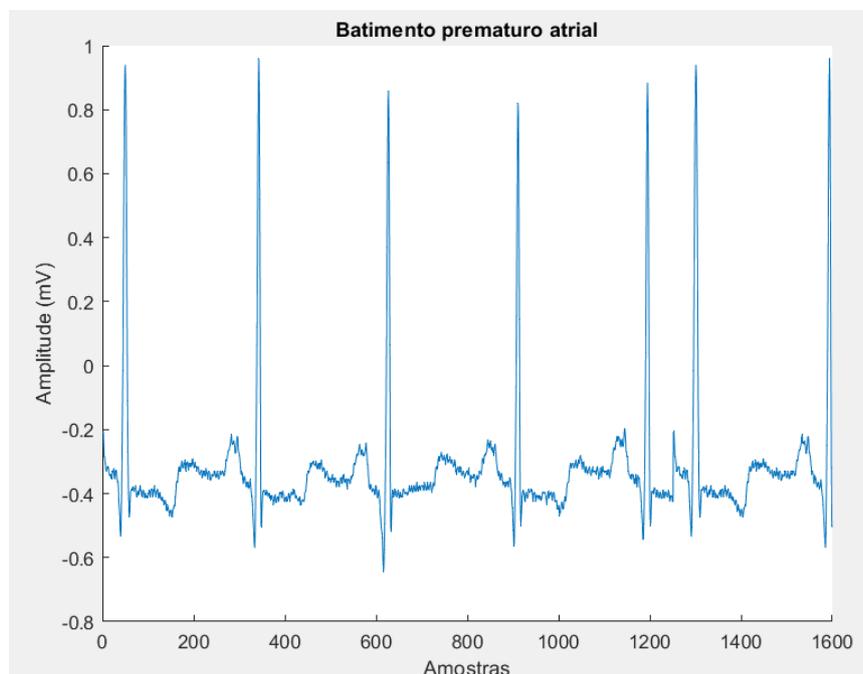


Figura 21. Formação de um subconjunto pertencente a classe batimento prematuro atrial (a).

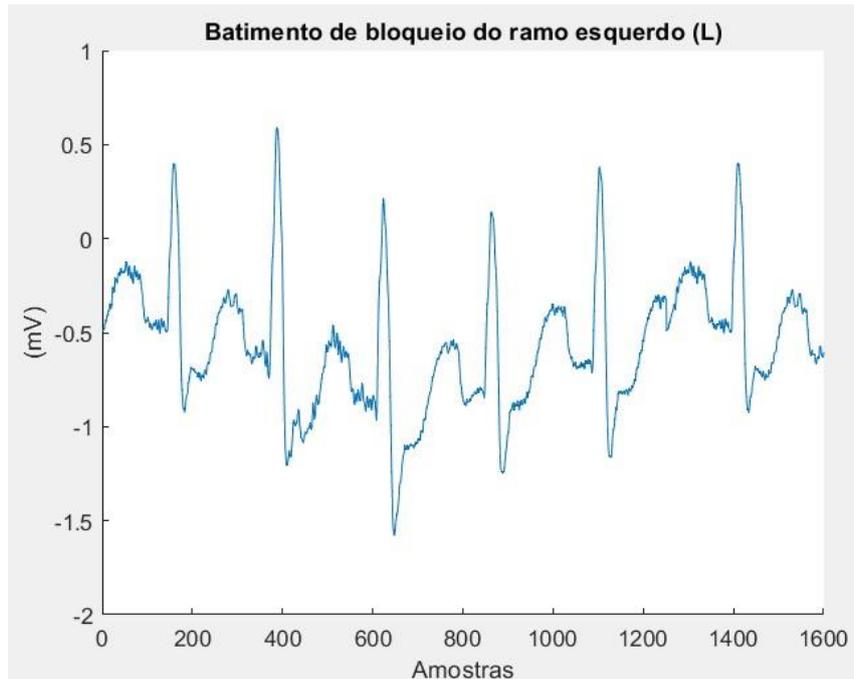


Figura 22. Formação de um subconjunto pertencente a classe batimento de bloqueio do ramo esquerdo (L).

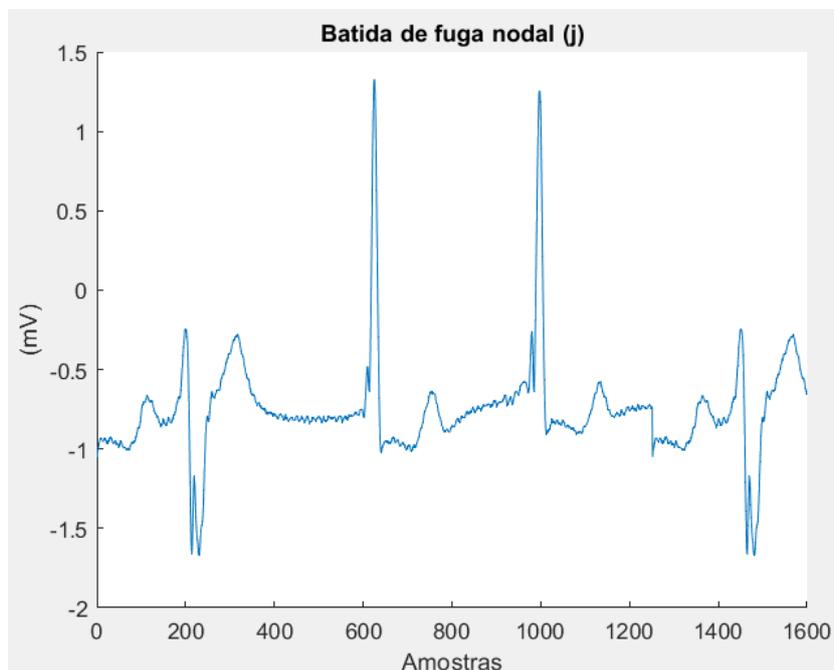


Figura 23. Formação de um subconjunto pertencente a classe batida de fuganodal (j).

#### 4.1.1 Reescalonamento do Sinal de ECG

Como apresentado nas Figuras 21,22 e 23, as amplitudes das amostras do sinal de ECG são expressas em tensão (mV). Considerando que trabalharemos com esse sinal (vetor 1D) convertido em uma matriz 2D (imagem em escala de cinza), faz-se necessário reescalonar as amplitudes dos sinais de ECG em uma escala equivalente aos níveis de cinza de uma imagem de 8 bits. Assim, é realizada a seguinte operação através da equação (8):

$$I_{esc} = \frac{(L-1)}{I_{max}-I_{min}} * (I - I_{min}) \quad (8)$$

Em que:

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

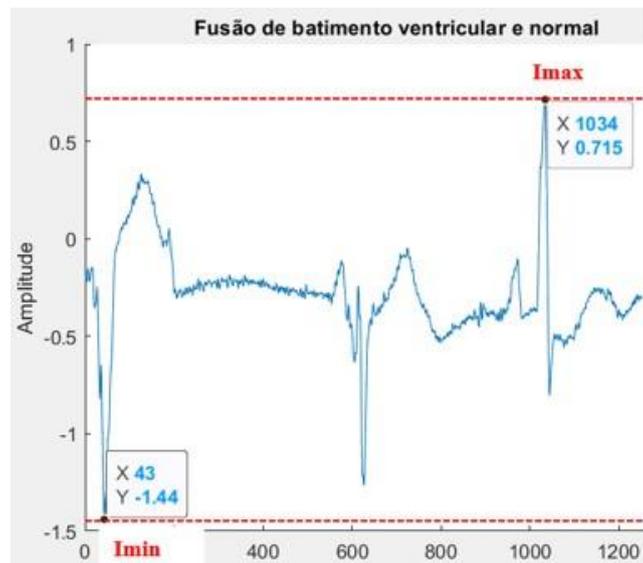
$I$  = Amplitude do Sinal de ECG a ser reescalonado;

$I_{min}$  = Valor mínimo da onda Q dos complexos QRS registrado em um subconjunto de valores do sinal de ECG;

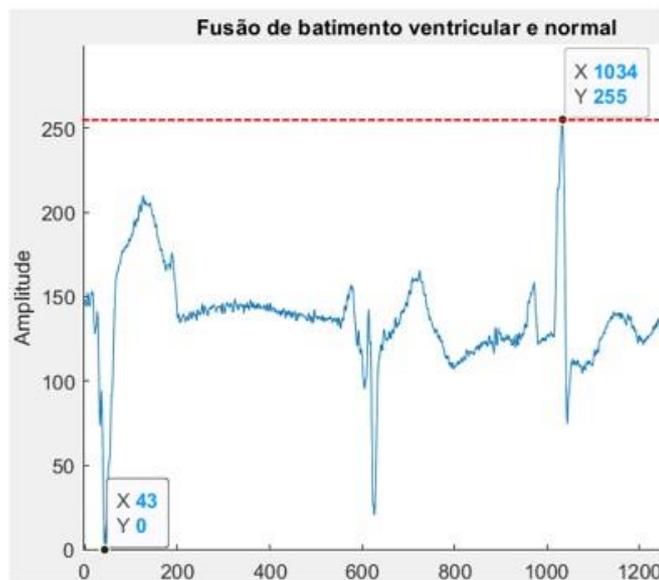
$I_{max}$  = Valor máximo da onda R dos complexos QRS registrados em um subconjunto de valores do sinal de ECG;

$I_{esc}$  = Amplitude do sinal de ECG reescalonado.

A Figura 24 mostra um subconjunto de valores de um sinal de ECG contendo a arritmia *fusão de batimento ventricular e normal*. Na Figura 24(a) mostra-se o sinal de ECG original em mV (eixo vertical), com os valores máximos e mínimos marcados (R e Q respectivamente). Na Figura 24(b) mostra-se o mesmo sinal de ECG reescalonado para a faixa [0,255] através da aplicação na equação (8) anteriormente mostrada.



(a)

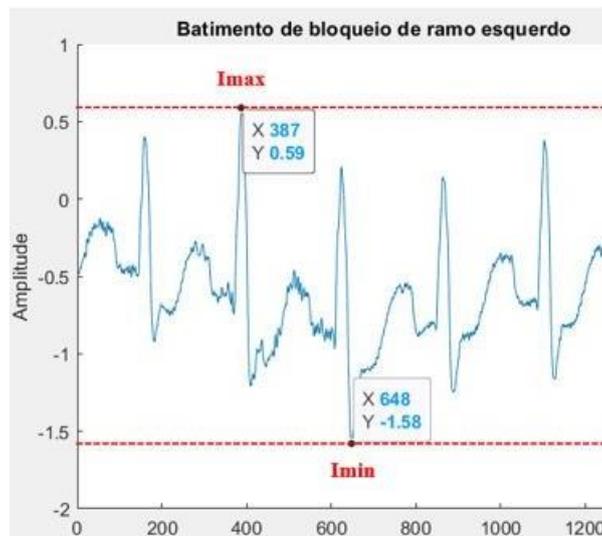


(b)

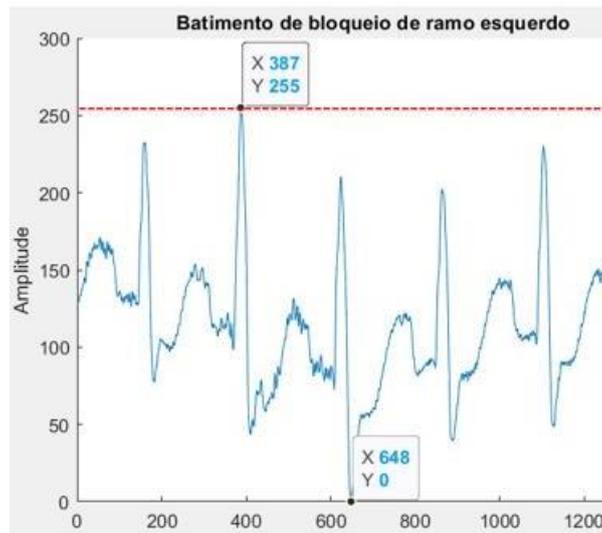
Figura 24. Processo de reescalamento aplicado ao subconjunto. (a) Sinal de ECG original com o eixo vertical expresso em mV. (b) Sinal de ECG reescalonado para a faixa [0,255].

Outros exemplos são apresentados nas Figuras 25 e 26.

Na Figura 25 temos um subconjunto de dados contendo um sinal de ECG com a arritmia *bloqueio do ramo esquerdo*. Na Figura 25(a), que mostra o sinal original, obtém-se os seguintes valores máximos e mínimos de tensão: valor máximo = 0,59mV, obtido em um pico de uma onda R de um complexo QRS, valor mínimo = -1.58mV, obtido em uma depressão de uma onda Q de um complexo QRS. Na Figura 25(b) mostra-se o mesmo sinal de ECG com os valores de tensão reescalonados para a faixa [0,255], utilizando a equação (8).



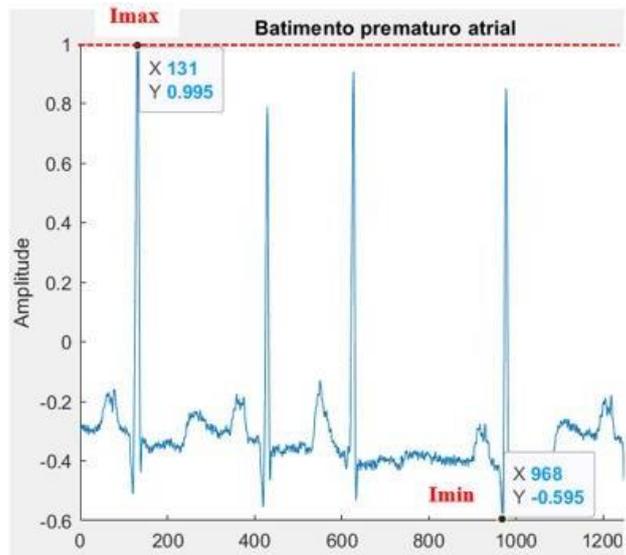
(a)



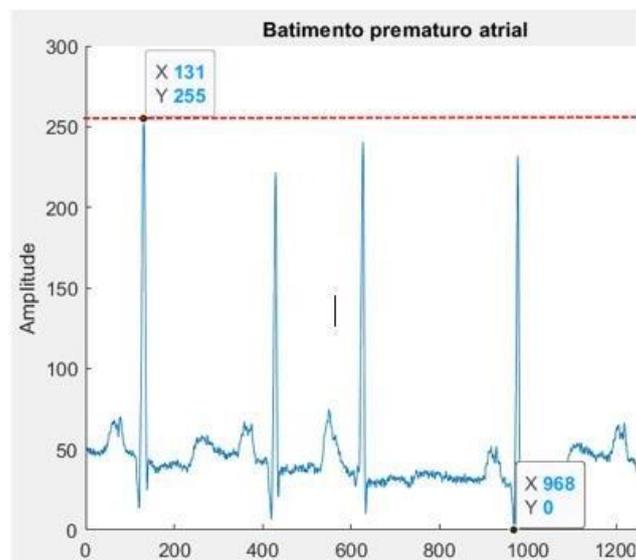
(b)

Figura 25. (a) Sinal original de ECG com arritmia bloqueio do ramo esquerdo, mostrando a determinação dos valores máximo e mínimo de tensão (b) Sinal de ECG com os valores de tensão reescalonados para a faixa [0,255].

Na Figura 26 temos um subconjunto de dados contendo um sinal de ECG com *batimento prematuro atrial*. Na Figura 26(a), que mostra o sinal original, obtém-se os seguintes valores máximos e mínimos de tensão: valor máximo = 0,995mV, obtido em um pico de uma onda R de um complexo QRS, valor mínimo = 0,595mV, obtido em uma depressão de uma dona Q de um complexo QRS. Na Figura 26(b) mostra-se o mesmo sinal de ECG os valores de tensão reescalados para a faixa [0,255], utilizando a equação (8).



(a)



(b)

Figura 26.(a) Sinal original de ECG com arritmia batimento prematuro atrial, mostrando a determinação dos valores máximo e mínimo de tensão (b) Sinal de ECG com os valores de tensão reescalados para a faixa [0,255].

#### 4.1.2 Conversão do Vetor 1D em Matriz 2D

Após o processo de reescalonamento dos valores contidos nos subconjuntos, a próxima etapa refere-se ao processo de conversão do sinal unidimensional de ECG contido em um subconjunto em uma matriz bidimensional, com dimensões 40x40. Cada elemento dessa matriz corresponde ao valor de uma amostra de um sinal de ECG e representará um pixel na imagem em níveis de cinza obtida a partir da matriz 2D anteriormente referida. Este processo de conversão 1D-2D é exemplificado na Figura 27.

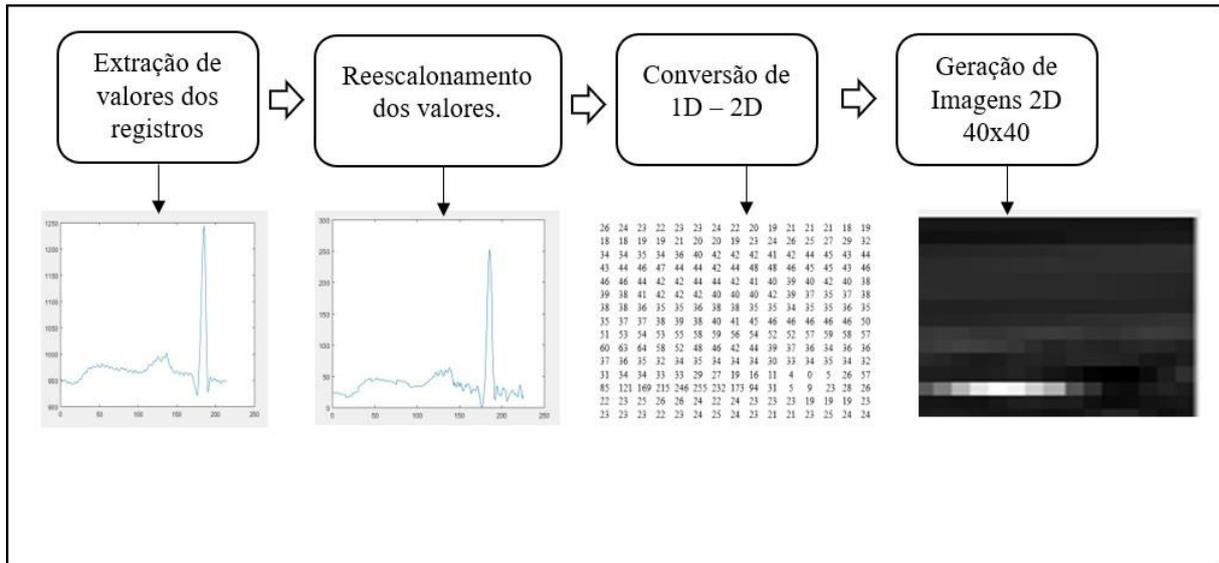


Figura 27. Processo de extração de um subconjunto a partir do sinal de ECG, reescalonamento, conversão 1D-2D e geração de imagens 2D.

A seguir na Figura 28(a) apresenta-se, para um subconjunto do ECG contendo uma arritmia batimento *prematureo nodal*, a matriz com dimensões de 40x40 com valores reescalados, enquanto na Figura 28(b) apresenta-se a imagem em escala de cinza de 8 bits obtida a partir dessa matriz.

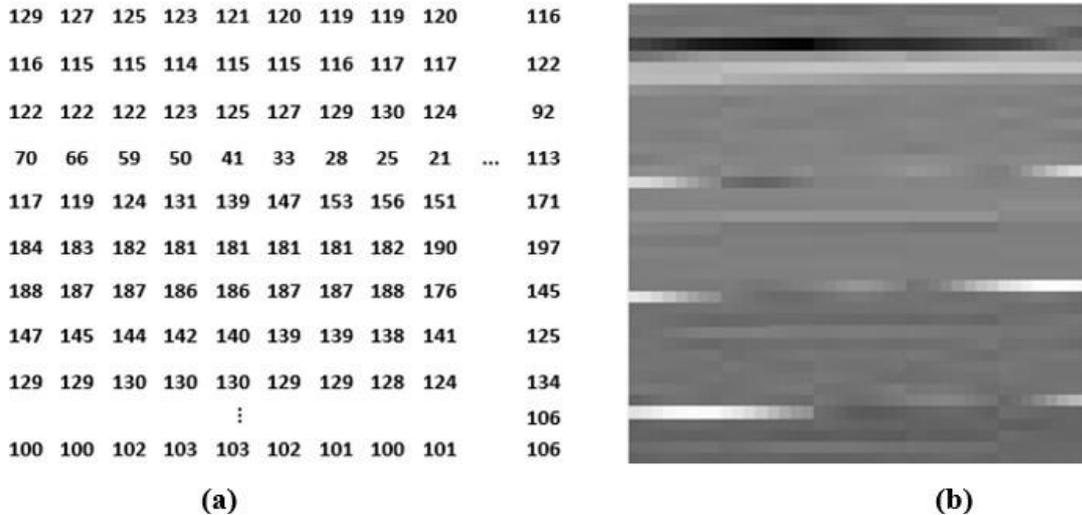


Figura 28. (a) subconjunto do ECG contendo uma arritmia batimento prematuro nodal, a matriz com dimensões de 40x40 com valores reescalados; (b) imagem em escala de cinza de 8 bits obtida a partir dessa matriz.

## 4.2 MÉTODOS

### 4.2.1 Aumento de dados

O número de amostras utilizadas nesse trabalho foi de 7402, divididas em 13 classes. Apesar da grande quantidade de amostras, verifica-se um grande desbalanceamento entre as classes. Por exemplo, a arritmia *batida de bloqueio de ramo direito* possui 7259 registros, enquanto a arritmia batimento *prematuro nodal* possui apenas 83 registros, e a arritmia *fuga atrial* possui menos registros, apenas 16. Com o objetivo de balancear as classes, nesse trabalho propõe-se um método para realizar o aumento de dados nas classes com poucos registros.

Cada gravação no banco de dados original possui cerca de 30 minutos. Nesse trabalho os subconjuntos de dados de ECG utilizados contêm uma janela com 1600 amostras. Para realizar o aumento dos dados, são criadas outras janelas a partir do deslocamento da janela original em diversas posições ao longo do sinal de ECG, em torno de uma arritmia. A quantidade de deslocamentos da janela varia de acordo com a classe utilizada, uma vez que por exemplo a arritmia *fuga atrial* que possui apenas 16 registros, ela necessita de uma maior quantidade de deslocamentos para que se obtenha uma maior quantidade de dados. Procura-se fazer o deslocamento da janela, mantendo-se sempre o complexo QRS em que se encontra a arritmia dentro dela. A localização do complexo QRS contendo a arritmia pode ocorrer nas seguintes direções: central, deslocamentos para à esquerda e deslocamentos para à direita. Na Figura 29, é mostrada apenas uma ilustração

dos deslocamentos das janelas geradas para um batimento contendo a arritmia  *fusão de batimento compassado e normal* que possui 982 registros retirados da base de dados. Nesta arritmia foram realizados apenas 3 deslocamentos ao todo.



Figura 29. Exemplo de aplicação da janela de captura. (a) janela com o complexo QRS contendo a arritmia bloqueio de ramo esquerdo centralizado; (b) janela com o complexo QRS contendo a arritmia bloqueio de ramo esquerdo à esquerda; (c) janela com o complexo QRS contendo a arritmia bloqueio de ramo esquerdo.

Observando a Figura 29, pode-se notar que o complexo QRS contendo a arritmia não é alterado e nem removido da janela de captura, o que difere é apenas a sua posição na janela. Na Figura 29, temos uma ilustração com 3 janelas. Como mencionado, pode haver mais deslocamentos, dependendo do número de eventos da arritmia contida na base de dados.

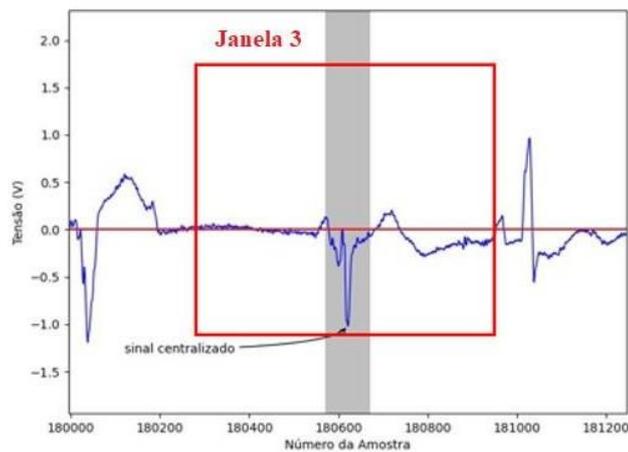
Na Figura 30 ilustramos outro exemplo de aumento de dados para um sinal contendo a arritmia  *batimento prematuro nodal*, extraído da base de dados. Nesta classe foram realizados 11 deslocamentos, porém são mostradas apenas 3. A faixa cinza, mostrada, refere-se à identificação do complexo QRS contendo a arritmia.



(a)



(b)



(c)

Figura 30. Exemplo da técnica utilizada para aumento de dados em um sinal contendo a arritmia batimento prematuro nodal. São apresentados apenas 3 deslocamentos de 11 realizados. (a) janela como complexo QRS contendo a arritmia batimento prematuro nodal à direita; (b) janela com o complexo QRS contendo a arritmia batimento prematuro nodal à esquerda; (c) janela com o complexo QRS contendo a arritmia batimento prematuro nodal centralizado.

Com os deslocamentos das janelas de captura dos subconjuntos do ECG mostrados, é possível obter novos subconjuntos de dados. O resultado do aumento de dados das classes selecionadas (classes com pequeno número de arritmias) é apresentado na Tabela 5.

Tabela 5. Aumento de dados do conjunto de treinamento das classes desbalanceadas através da técnica de aumento de dados proposta

<b>Símbolo</b>	<b>Arritmia</b>	<b>Original</b>	<b>Treinamento 70% (sem aumento)</b>	<b>Treinamento (com aumento)</b>	
<b>1</b>	A	Batimento atrial prematureo	150	100	700
<b>2</b>	J	Batimento prematureo nodal	83	58	638
<b>3</b>	E	Batida de fuga atrial	16	11	121
<b>4</b>	J	Batida de fuga nodal	229	160	800
<b>5</b>	E	Batida de escape ventricular	106	74	740
<b>6</b>	Q	Batida inclassificável	33	23	253
		<b>Total</b>	617		3252

Das 13 classes de arritmias, aplicou-se a técnica de aumento de dados em 6 classes, (conforme mostra a Tabela 5). As demais classes possuíam um número de subconjuntos suficiente para a realização de treinamento e validação.

#### 4.2.2 Arquitetura da rede CNN proposta

Neste trabalho, a classificação das 13 classes de arritmias cardíacas é realizada através de uma rede convolucional (vide figura 31). A entrada da rede recebe imagens em escala de cinza de dimensão 40x40 pixels. As características são extraídas através de blocos conectados em sequência e constituídos das camadas *conv-batch-relu*, que representam, respectivamente as camadas de convolução, normalização em lote e função de ativação. A camada de convolução possui dimensão igual a 96x96 pixels, com filtros de tamanhos 3x3 aplicados com passo 1. Em sequência, temos 2 camadas inteiramente conectadas (*FCLayer*). Cada camada possui um total de 200 neurônios. Entre essas camadas inteiramente conectadas, com o objetivo de melhorar a generalização da rede, foi inserida uma camada de *Dropout*. Na última camada temos a *Softmax*, que contém 13 neurônios na saída e é responsável pela classificação

Com o objetivo de alcançar uma arquitetura de alto desempenho, foi conduzido um estudo de ablação. O estudo consistiu em repetir o processo de treinamento da arquitetura variando o número de blocos *conv-batch-relu*. Foram testadas três quantidades de blocos: 3, 5 e 7. A Figura 31 apresenta as representações visuais das três arquiteturas. Em 31(a) temos a primeira arquitetura proposta (CNN2D-1) contendo 3 blocos *conv-batch-relu*, 31(b) a segunda arquitetura proposta contendo 5 blocos (CNN2D-2), e em 31(c) a terceira arquitetura proposta contendo 7 blocos (CNN2D-3). É importante destacar que os parâmetros anteriormente mencionados foram determinados através de diversos testes.

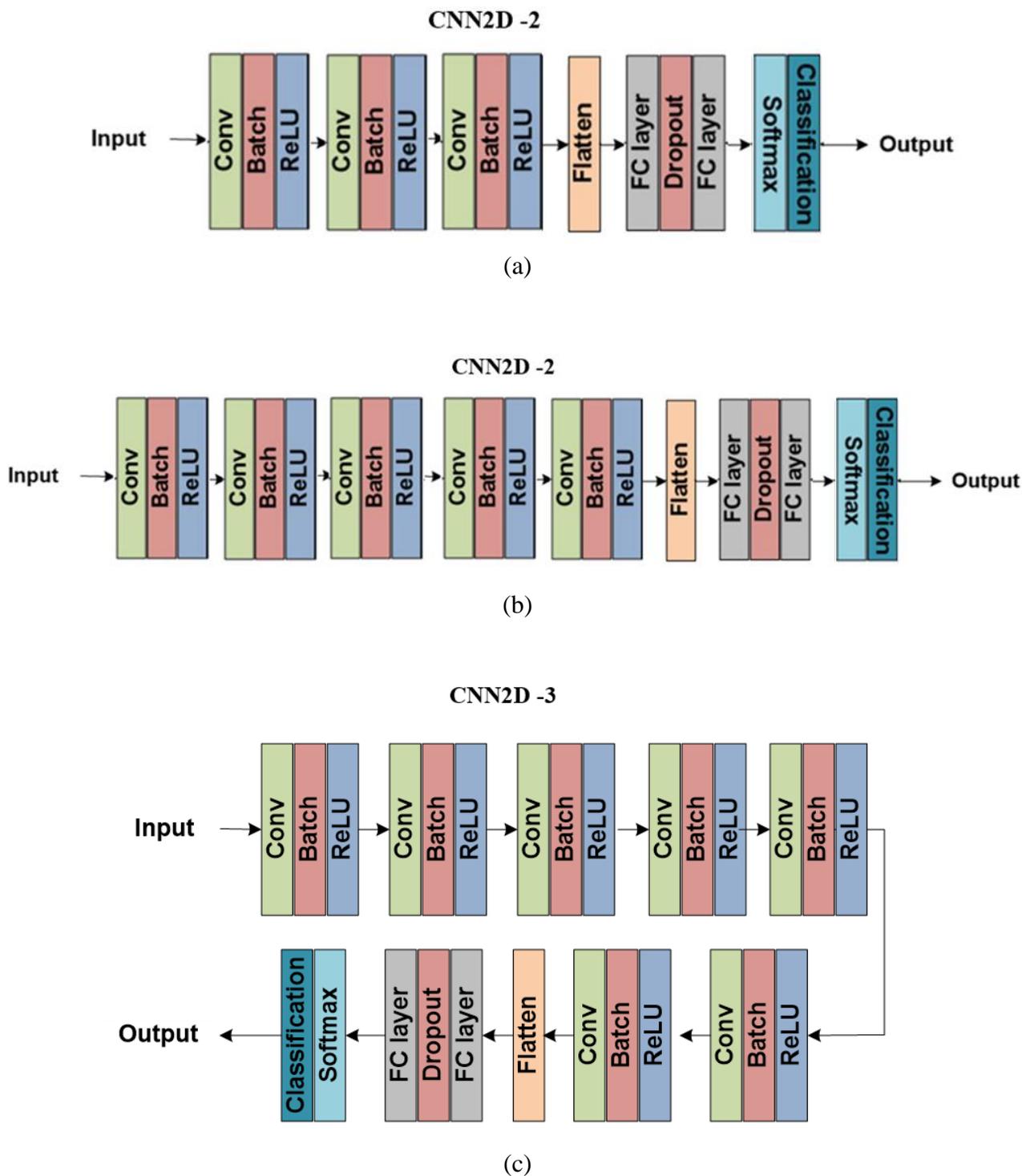


Figura 31. Arquiteturas avaliadas neste estudo (a) CNN2D-1 contendo 3 blocos *conv-batch-reLU*. (b) CNN2D-2 contendo 5 blocos e (c) CNN2D-3 contendo 7 blocos.

Conforme mostrado na Figura 31, as arquiteturas diferem nas quantidades de blocos *conv-batch-reLU*, entretanto a forma como esses blocos são configurados não mudam. A camada de convolução possui dimensão igual a 96x96, com filtros de tamanhos 3x3 aplicados com passo 1.

### 4.2.3 Treinamento da rede

O processo de treinamento das três arquiteturas foi realizado conforme o fluxograma apresentado na Figura 32. A divisão do conjunto de dados foi a seguinte: 70% das imagens foram utilizadas para a formação do conjunto de treinamento, enquanto 15% foram destinados a à formação do conjunto de validação e 15%, ao conjunto de teste.



Figura 32. Fluxograma de como foi realizado o treinamento das três arquiteturas propostas.

De acordo com o fluxograma apresentado, cada uma das três arquiteturas propostas passou por dois ciclos de treinamento, validação e testes. Inicialmente, utilizou-se o conjunto de dados original para o treinamento, seguido por um segundo ciclo de treinamento, no qual as amostras foram aumentadas nas seis classes selecionadas. Esse procedimento permitiu uma avaliação abrangente do desempenho de cada arquitetura em duas configurações distintas de dados. Durante as simulações também foram realizados ajustes em alguns hiper parâmetros: 1) taxa de aprendizado: valor inicial de 0,01 com fator de decaimento de 0,75; 2) número de épocas do treinamento: 100 épocas; 3) tamanho do minilote: 512; 4) Taxa de regularização  $L_2$ : 0,1; 5) Fator de *dropout*: 0,5.

O ambiente utilizado nas simulações das redes CNN foi o *MATLAB 2021*. O *hardware* utilizado foi um computador com as especificações: processador Intel(R) Core(TM) i7-12700H 2.30 GHz; memória RAM 32 GB; Placa de vídeo NVIDIA 16GB RTX.

### 4.2.4 Métricas de Avaliação

Para avaliar o desempenho da arquitetura proposta nesse trabalho, foram utilizadas 5 métricas. Para a definição dessas métricas, são necessários os seguintes conceitos:

1) VP: verdadeiro positivo, quando uma amostra contendo uma arritmia é detectada corretamente pelo classificador;

- 2) VN: verdadeiro negativo, quando uma amostra normal (que não contém uma arritmia) é detectada corretamente pelo classificador;
- 3) FP: falso positivo, quando uma amostra normal (que não contém uma arritmia) é detectada incorretamente pelo classificador;
- 4) FN: falso negativo, quando uma amostra contendo uma determinada arritmia é classificada incorretamente pelo classificador.

Com base nesses conceitos apresenta-se as seguintes métricas utilizadas:

- Acurácia: determina com que frequência o classificador classifica corretamente.

$$Acurácia = \frac{VP+TN}{VP+FP+TP+TN} \quad (9)$$

- Precisão: determina a fração de amostras classificadas como positivas que eram realmente positivas.

$$Precisão = \frac{VP}{VP+FP} \quad (10)$$

- Sensibilidade (*Recall*): determina com que frequência o classificador classifica corretamente os exemplos da classe positiva. Quanto maior o valor da sensibilidade, menos FN são detectados pelo classificador.

$$Sensibilidade = \frac{VP}{VP+FN} \quad (11)$$

- Especificidade: determina com que frequência o classificador classifica corretamente os exemplos da classe negativa. Quanto maior o valor da especificidade, menos FP são detectados pelo classificador.

$$Especificidade = \frac{VN}{VN+FP} \quad (12)$$

- *F1-score*: é calculado pela combinação das métricas de precisão e sensibilidade, indica a qualidade geral do modelo.

$$F1 - score = 2 * \frac{(Precisão * Sensibilidade)}{Precisão + Sensibilidade} \quad (13)$$

#### 4.2.5 Técnica de visualização de ativação

Após a classificação das arritmias cardíacas, foi proposto neste trabalho destacar, de forma visual, o trecho do sinal de ECG onde a arritmia está contida. Para isso foi utilizada a técnica de mapeamento de ativação através do gradiente (do inglês *Gradient-weighted Class Activation Mapping*) também conhecida como *Grad-CAM*.

O *GradCAM* é uma técnica de interpretação de modelos de redes neurais convolucionais que ajuda a entender quais partes da imagem contribuíram mais para a decisão de classificação feita pela rede.

A ideia por trás do *GradCAM* é gerar um mapa de calor que destaca as regiões da imagem que mais influenciaram a predição da classe. Isso é feito calculando os gradientes da saída da última camada convolucional em relação à classe de interesse e, em seguida, combinando esses gradientes com as ativações da mesma camada (SELVARAJU *et al.*, 2017). Isso cria um mapa de ativação ponderado que mostra quais partes da imagem são mais importantes para a decisão da classe. Os passos para cálculo do gradiente são demonstrados a seguir.

- Seja  $L$  a última camada convolucional da rede neural.
- Seja  $F_k$  a ativação da  $k$ -ésima unidade (*feature map*) na camada  $L$ .
- Seja  $Y_c$  a probabilidade de classificação da classe  $c$ .

O gradiente da classe  $c$  é calculado em relação à ativação da camada  $L$  através da equação (14).

$$\frac{\partial Y_c}{\partial F_k} \quad (14)$$

O peso de importância ( $A_k^{(c)}$ ) de cada *feature map*  $F_k$  para classe  $c$  é calculado usando a média ponderada conforme equação (15).

$$A_k^{(c)} = \frac{1}{Z} \sum_i \sum_j \frac{\partial Y_c}{\partial F_k(i,j)} \quad (15)$$

em que  $Z$  é uma constante de normalização.

Após isso é aplicada uma função de ativação *ReLU* para obter as ativações ponderadas, equação (16).

$$L_k^{(c)} = \max(0, A_k^{(c)}) \quad (16)$$

Finalmente é calculado o gradiente combinando as ativações ponderadas com o *features maps* originais através da equação (17).

$$GradCAM_c = \sum_k L_k^{(c)} F_k \quad (17)$$

O *Grad-CAM* acessa a saída de uma camada convolucional e, a partir das informações presentes na saída da mesma, gera um mapa de pontuação (do inglês *Score Map*). Essas informações extraídas como mencionado anteriormente, representam o impacto das características extraídas na saída da camada convolutiva sobre a classificação da CNN. Neste trabalho, o mapa de pontuação baseado nessas características, gerou valores no intervalo [0 1] e possuem a mesma dimensão que as imagens de entrada, ou seja, os mapas gerados possuem dimensão de 40x40. Na Figura 33 mostra-se um exemplo de um mapa de pontuação.

0.2893	0.2941	0.2955	0.2921	0.2842
0.1988	0.2351	0.2543	0.2520	0.2337
0.1883	0.2146	0.2319	0.2318	0.2154
0.3242	0.3367	0.3539	0.3659	0.3692
0.5016	0.5234	0.5500	0.5743	0.5950
0.4644	0.4692	0.4778	0.4825	0.4783
0.3952	0.3863	0.3812	0.3742	0.3610
0.3506	0.3543	0.3560	0.3563	0.3566
0.5040	0.5128	0.5154	0.5147	0.5137
0.5759	0.5774	0.5699	0.5588	0.5515
0.4675	0.4722	0.4709	0.4674	0.4673
0.2837	0.3056	0.3302	0.3551	0.3746
0.1229	0.1301	0.1398	0.1535	0.1734
0.2310	0.2696	0.3022	0.3081	0.2956

Figura 33. Trecho de um mapa de pontuação gerado pelo GradCAM.

Por fim, as pontuações contidas no mapa acima podem ser associadas a imagem de entrada através de um mapa de calor. Na Figura 34, mostra-se uma imagem aplicada à entrada da rede CNN2D com seu respectivo mapa de calor.

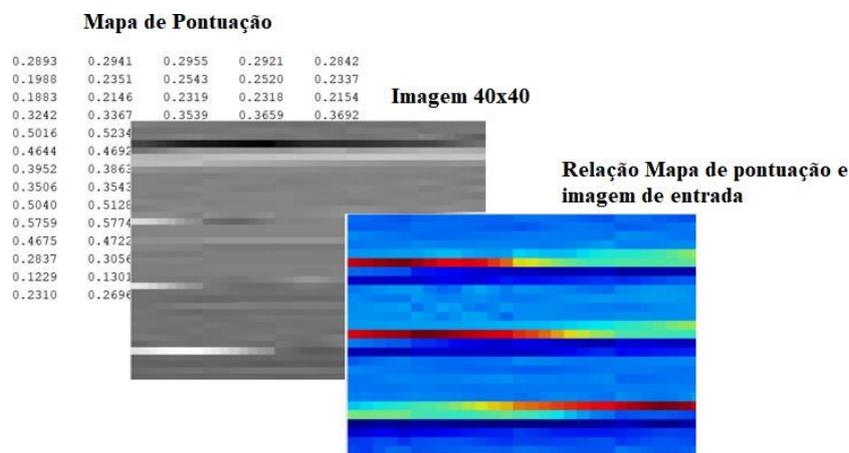


Figura 34. Exemplo de imagem aplicada à entrada da rede CNN2D, com seu respectivo mapa de calor gerado pelo método Grad-CAM.

As cores contidas no mapa de calor representam o quanto foi significante aquela região para decisão da rede. As cores mais frias são representadas por tonalidades próximas ao azul, e representam as áreas com baixas probabilidades de conter uma arritmia cardíaca, enquanto as cores mais quentes, representadas pelas tonalidades próximas ao vermelho, representam as áreas com alta probabilidade de conter uma arritmia cardíaca.

A proposta neste trabalho, então, foi utilizar técnicas de visualização de ativação *Grad-CAM* para mostrar no próprio sinal de ECG a área que contém a arritmia. Essa representação é realizada em três etapas, conforme ilustrado na Figura 36.

Na primeira etapa, mostrada na Figura 35(a), é gerado uma matriz com dimensões de 40x40, contendo as pontuações geradas pelo método *Grad-CAM*. Na Figura 35(a) é também mostrada uma imagem de intensidade correspondendo a essa matriz. Na segunda etapa, mostrada na Figura 35(b), a matriz 2D gerada na etapa anterior, é convertida em um vetor de 1600 coordenadas, correspondendo às coordenadas de um subconjunto do ECG. Finalmente, na Figura 35(c), destaca-se, no sinal de ECG, as regiões do vetor determinado na etapa anterior com valor maior ou igual a um limiar de 0,8. Este valor de limiar foi obtido através de diversos testes (será explicado na sessão de resultados), significando que, onde há uma região com arritmia, certamente a pontuação é igual ou superior a 0,8 (regiões dentro dos retângulos vermelhos).



## 5. RESULTADOS E DISCUSSÕES

Neste capítulo apresentamos os resultados obtidos pelo método proposto nesse trabalho para classificação e detecção das arritmias cardíacas. Esses resultados são os seguintes: comparação entre as curvas de convergência da acurácia, durante o treinamento, validação e testes das três arquiteturas de rede CNN proposta através do uso de dois diferentes otimizadores SGDM e ADAM e os resultados das classificações das 13 classes de arritmia cardíacas antes e após o aumento de dados.

### 5.1 Método de Otimização ADAM

O primeiro otimizador utilizado foi o ADAM. Foram conduzidas 100 épocas nos três processos de treinamento executados. Na Tabela 6 são apresentados os resultados referentes a três arquiteturas propostas, antes e depois da aplicação do aumento de dados. Também são apresentados os resultados das acurácias médias (média aritmética das acurácias das 13 classes) e das acurácias globais (quantidade total de acertos dividida pelo total de amostras).

Tabela 6. Resultados obtidos para as três arquiteturas propostas utilizando o método de otimização ADAM.

Classe	CNN2D-1		CNN2D-2		CNN2D-3	
	<u>Aumento de dados</u>					
	Sem	Com	Sem	Com	Sem	Com
<b>E</b>	6,70%	73,30%	80,00%	93,30%	0,00%	80,00%
<b>F</b>	85,80%	75,00%	95,80%	85,00%	89,20%	76,70%
<b>J</b>	0,00%	64,30%	0,00%	92,90%	0,00%	92,90%
<b>L</b>	72,00%	90,70%	84,70%	91,30%	75,30%	88,00%
<b>N</b>	98,70%	100,00%	98,70%	98,70%	98,70%	98,70%
<b>P</b>	96,70%	94,70%	96,70%	97,30%	95,30%	98,00%
<b>Q</b>	0,00%	20,00%	0,00%	20,00%	0,00%	100,00%
<b>R</b>	98,70%	98,70%	98,70%	100,00%	100,00%	98,70%
<b>V</b>	84,70%	84,70%	94,00%	78,70%	97,30%	72,00%
<b>a</b>	88,00%	100,00%	80,00%	96,00%	84,00%	96,00%
<b>e</b>	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
<b>f</b>	41,50%	53,10%	64,60%	63,30%	51,70%	50,30%
<b>j</b>	0,00%	55,90%	0,00%	58,80%	2,90%	55,90%
<b>Acurácia Média</b>	51,75%	71,44%	61,02%	75,02%	53,42%	<b>77,48%</b>
<b>Acurácia Global</b>	77,54%	84,10%	85,44%	86,79%	81,31%	<b>82,84%</b>

Ao analisar a Tabela 6, é possível observar o impacto do aumento de dados no desempenho das arquiteturas. Inicialmente, na primeira arquitetura com 3 blocos, sem aumento de dados, a acurácia média era de 51,75%, mas após a aplicação do aumento de dados, esse valor aumentou para 71,44%. Na segunda arquitetura com 5 blocos, a acurácia média inicial sem aumento de dados era de 61,02%, enquanto com aumento de dados, esse valor aumentou para 75,02%. Na terceira arquitetura com 7 blocos, a acurácia média inicial sem aumento de dados era de 53,42%, e com aumento de dados, esse valor melhorou para 77,48%.

No que diz respeito à acurácia global, os resultados foram os seguintes: na primeira arquitetura sem aumento de dados, a acurácia global era de 77,54%, e após o aumento de dados, esse valor aumentou para 84,10%. Na segunda arquitetura sem aumento de dados, a acurácia global era de 85,44%, e após o aumento de dados, esse valor aumentou para 86,79%. Na terceira arquitetura sem aumento de dados, a acurácia global era de 81,31%, e após o aumento de dados, esse valor aumentou para 82,84%.

Embora o aumento de dados tenha tido um efeito positivo nos resultados de acurácia média e acurácia global durante o treinamento das três arquiteturas, ao analisarmos o desempenho individual de algumas classes de arritmia, como é o caso da classe de batida de fuga atrial (e), observamos que, nas três arquiteturas, com ou sem aumento de dados utilizando o otimizador ADAM, a rede não conseguiu aprender as características específicas dessa classe. Similarmente, ao considerarmos outra classe, como a de batidas não classificadas (Q), as arquiteturas 1 e 2 também não conseguiram aprender suas características.

Diante dos resultados obtidos nos testes das três arquiteturas, fica evidente que o método de otimização ADAM, quando associado à arquitetura proposta e seus hiperparâmetros, não gerou resultados satisfatórios, especialmente no que diz respeito à identificação de características específicas em determinadas classes de arritmia, a arquitetura que apresentou o melhor desempenho foi a arquitetura CNN2D-3.

## **5.2 Método de Otimização SGDM**

O segundo otimizador utilizado foi o SGDM. Foram conduzidas 100 épocas nos três processos de treinamento executados. Na Tabela 7 a seguir, são apresentados os resultados referentes a três arquiteturas propostas, antes e depois da aplicação do aumento de dados, também são apresentados os resultados das acurácias médias (média aritmética

das acurácias das 13 classes) e das acurácias globais (quantidade total de acertos dividida pelo total de amostras).

Tabela 7. Resultados obtidos para as três arquiteturas propostas utilizando o método de otimização SGDM.

Classe	CNN2D-1		CNN2D-2 Aumento de dados		CNN2D-3	
	Sem	Com	Sem	Com	Sem	Com
<b>E</b>	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%
<b>F</b>	99,20%	95,80%	100,00%	97,50%	98,30%	91,70%
<b>J</b>	0,00%	92,90%	0,00%	92,90%	7,10%	85,70%
<b>L</b>	98,00%	90,70%	90,70%	96,00%	97,30%	98,70%
<b>N</b>	100,00%	98,70%	98,70%	100,00%	100,00%	100,00%
<b>P</b>	97,30%	97,30%	96,70%	97,30%	97,30%	97,30%
<b>Q</b>	0,00%	0,00%	0,00%	100,00%	0,00%	100,00%
<b>R</b>	98,70%	98,70%	100,00%	98,70%	100,00%	100,00%
<b>V</b>	82,00%	68,00%	86,70%	76,00%	95,30%	81,30%
<b>a</b>	96,00%	100,00%	100,00%	100,00%	96,00%	100,00%
<b>e</b>	0,00%	0,00%	0,00%	100,00%	0,00%	100,00%
<b>f</b>	100,00%	93,20%	100,00%	100,00%	100,00%	99,30%
<b>j</b>	8,80%	91,20%	0,00%	100,00%	5,90%	100,00%
<b>Acurácia Média</b>	67,69%	78,96%	67,14%	89,10%	96,00%	<b>96,46%</b>
<b>Acurácia Global</b>	91,80%	91,30%	91,30%	95,30%	93,40%	<b>95,80%</b>

Iniciando a análise pela primeira arquitetura, composta por 3 blocos de extração de características, observa-se que sem o aumento de dados a acurácia global obtida foi de 91,80% e a acurácia média foi de 67,69%, a seguir na Figura 36, temos a curva de convergência da acurácia ao longo do treinamento, para os conjuntos de validação e teste.

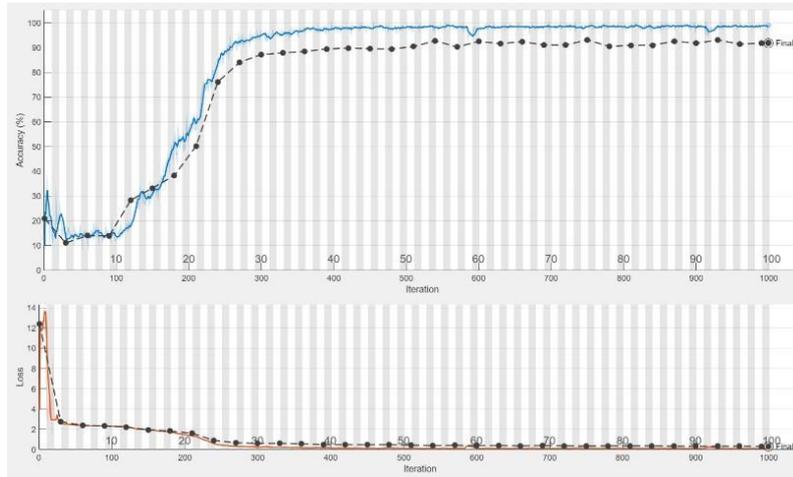


Figura 36. Curva de convergência da acurácia ao longo do treinamento, para os conjuntos de validação e teste, utilizando-se o método SGDM com 3 blocos de extração de características sem aumento de dados.

Ao analisarmos as curvas na Figura 36, são notadas oscilações apenas nas 25 primeiras épocas do treinamento. É importante ressaltar que a rede se estabiliza a partir da 30ª época. Na Figura 37 é mostrada a matriz de confusão obtida no conjunto de testes.

**Matriz de Confusão**

	M	F	J	L	N	P	Q	R	V	a	e	l	L	
M	15 1.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
F	0 0.0%	119 10.7%	0 0.0%	3 0.3%	0 0.0%	2 0.2%	0 0.0%	0 0.0%	1 0.1%	1 0.1%	0 0.0%	0 0.0%	0 0.0%	94.4% 5.6%
J	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
L	0 0.0%	0 0.0%	0 0.0%	147 13.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
N	0 0.0%	0 0.0%	1 0.1%	0 0.0%	150 13.5%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 0.1%	0 0.0%	1 0.1%	98.0% 2.0%
P	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	146 13.1%	0 0.0%	0 0.0%	14 1.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	91.2% 8.8%
Q	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
R	0 0.0%	0 0.0%	1 0.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	148 13.3%	1 0.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	98.7% 1.3%
V	0 0.0%	0 0.0%	3 0.3%	0 0.0%	0 0.0%	0 0.0%	2 0.2%	0 0.0%	123 11.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	96.1% 3.9%
a	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	24 2.2%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
e	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
l	0 0.0%	0 0.0%	8 0.7%	0 0.0%	0 0.0%	2 0.2%	3 0.3%	2 0.2%	11 1.0%	0 0.0%	2 0.2%	147 13.2%	30 2.7%	71.7% 28.3%
L	0 0.0%	1 0.1%	1 0.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	3 0.3%	60.0% 40.0%
	100% 0.0%	99.2% 0.8%	0.0% 100%	98.0% 2.0%	100% 0.0%	97.3% 2.7%	0.0% 100%	98.7% 1.3%	82.0% 18.0%	96.0% 4.0%	0.0% 100%	100% 0.0%	8.8% 91.2%	91.8% 8.2%
		←	→	↙	↘	↖	↗	↕	↔	↔	↕	↕	↕	
		<b>Classe Predita</b>												

Figura 37. Matriz de confusão para o conjunto de teste, utilizando-se o método SGDM com 3 blocos de extração de características sem aumento de dados.

Ao analisarmos a matriz de confusão apresentada na Figura 37 é possível verificar que esta arquitetura já apresenta um resultado bom para algumas classes, como é o caso da classe Batida de escape ventricular (E), Batida normal (N) e Fusão de batimento compassado e normal (f) que obtiveram acertos de 100% no conjunto de testes, porém ao analisarmos as classes Batimento prematuro nodal (J), Batimento não classificado (Q) e Batida de fuga atrial (e) notamos a dificuldade da arquitetura em aprender as características resultando em nenhum acerto no conjunto de testes.

Após o processo de treinamento utilizando o aumento de dados, o valor da acurácia global reduziu para 91,30% enquanto a acurácia média elevou-se para 78,96%. A seguir na Figura 38 é mostrada a nova curva de convergência da acurácia ao longo do treinamento, para os conjuntos de validação e teste.

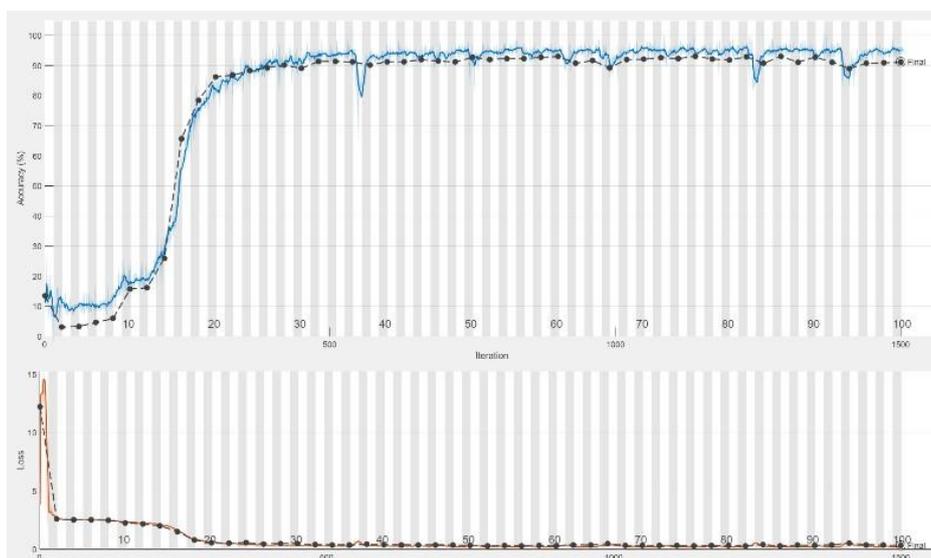


Figura 38. Curva de convergência da acurácia ao longo do treinamento utilizando o aumento de dados, para os conjuntos de validação e teste, utilizando-se o métodoSGDM com 3 blocos de extração de características após o aumento de dados.

Ao analisarmos as curvas na Figura 38, são notadas pequenas oscilações próximas a 40<sup>a</sup>, 80<sup>a</sup> e 95<sup>a</sup> épocas causadas pelo acréscimo de novas imagens durante o treinamento. Na Figura 39 é mostrada a matriz de confusão obtida no conjunto de testes.

### Matriz de Confusão

Classe verdadeira	E	15 1.3%	1 0.1%	0 0.0%	5 0.4%	0 0.0%	0 0.0%	0 0.0%	2 0.2%	3 0.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	57.7%
	F	0 0.0%	115 10.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100%
	J	0 0.0%	0 0.0%	13 1.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 0.1%	0 0.0%	2 0.2%	0 0.0%	1 0.1%	76.5%
	L	0 0.0%	0 0.0%	0 0.0%	136 12.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100%
	N	0 0.0%	0 0.0%	0 0.0%	0 0.0%	148 13.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 0.1%	99.3%
	P	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	146 13.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	8 0.7%	0 0.0%	94.8%
	Q	0 0.0%	2 0.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 0.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 0.1%	0.0%
	R	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	148 13.3%	1 0.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	99.3%
	V	0 0.0%	0 0.0%	0 0.0%	9 0.8%	0 0.0%	0 0.0%	1 0.1%	0 0.0%	102 9.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	91.1%
	a	0 0.0%	2 0.2%	1 0.1%	0 0.0%	2 0.2%	4 0.4%	1 0.1%	0 0.0%	39 3.5%	25 2.2%	0 0.0%	2 0.2%	0 0.0%	32.9%
	e	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN%
	f	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	3 0.3%	0 0.0%	3 0.3%	0 0.0%	1 0.1%	137 12.3%	0 0.0%	95.1%
	j	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	31 2.8%	100%
		100%	95.8%	92.9%	90.7%	98.7%	97.3%	0.0%	98.7%	68.0%	100%	0.0%	93.2%	91.2%	91.3%
		0.0%	4.2%	7.1%	9.3%	1.3%	2.7%	100%	1.3%	32.0%	0.0%	100%	6.8%	8.8%	8.7%

Figura 39. Matriz de confusão para o conjunto de teste, utilizando-se o método SGDM com 3 blocos de extração de características após aumento de dados.

Ao analisarmos a nova matriz de confusão apresentada na Figura 39 é possível verificar que a rede apresentou melhorias em algumas classes como é o caso da classe Batimento prematuro nodal (J) que agora apresentou um acerto de 92,9% (sem aumento apresentou 0% de acerto), outro exemplo é a classe Batida de fuga nodal (j) que agora apresentou um acerto de 91,2% (sem aumento apresentou 8,8% de acerto), entretanto ao analisarmos as classes Batimento não classificado (Q) e Batida de fuga atrial (e), notamos ainda a dificuldade da arquitetura em aprender as características resultando novamente em nenhum acerto no conjunto de testes.

A segunda arquitetura, composta por 5 blocos de extração de características, observa-se que sem o aumento de dados o valor obtido da acurácia global foi 91,30% e a acurácia média foi de 67,14%, a seguir na Figura na Figura 40, temos a curva de convergência da acurácia ao longo do treinamento, para os conjuntos de validação e teste.

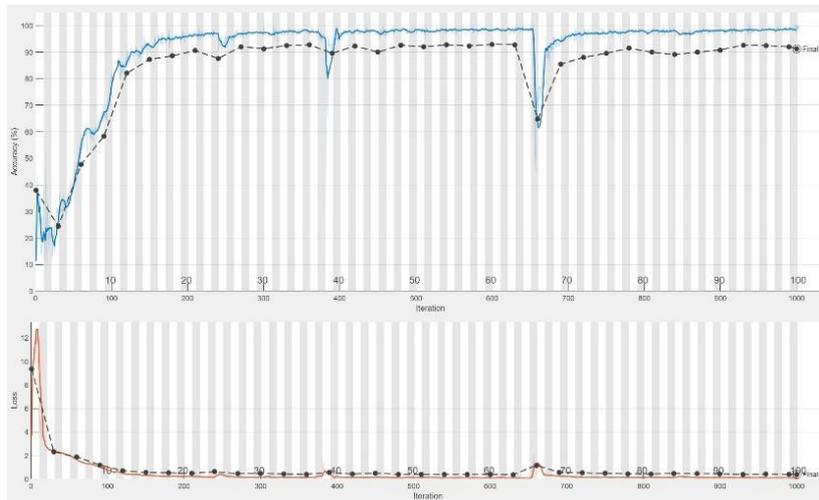


Figura 40. Curva de convergência da acurácia ao longo do treinamento, para os conjuntos de validação e teste, utilizando-se o método SGDM com 5 blocos de extração de características sem o aumento de dados.

Ao analisarmos as curvas na Figura 40, são notadas oscilações significativas nas 40<sup>a</sup> e 70<sup>a</sup> épocas, a rede se estabiliza a partir da 75<sup>a</sup> época. A Figura 41 mostra a matriz de confusão obtida no conjunto de testes.

**Matriz de Confusão**

<b>Classe verdadeira</b>	E	15 1.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%	
	F	0 0.0%	120 10.8%	0 0.0%	3 0.3%	0 0.0%	3 0.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	95.2% 4.8%
	J	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
	L	0 0.0%	0 0.0%	0 0.0%	136 12.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	N	0 0.0%	0 0.0%	6 0.5%	0 0.0%	148 13.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 0.1%	0 0.0%	2 0.2%	94.3% 5.7%
	P	0 0.0%	0 0.0%	0 0.0%	0 0.0%	2 0.2%	145 13.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	98.6% 1.4%
	Q	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
	R	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	150 13.5%	1 0.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	99.3% 0.7%
	V	0 0.0%	0 0.0%	2 0.2%	11 1.0%	0 0.0%	0 0.0%	2 0.2%	0 0.0%	130 11.7%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	89.7% 10.3%
	a	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	25 2.2%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	e	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
	f	0 0.0%	0 0.0%	6 0.5%	0 0.0%	0 0.0%	2 0.2%	3 0.3%	0 0.0%	19 1.7%	0 0.0%	2 0.2%	147 13.2%	32 2.9%	69.7% 30.3%
	L	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
			100% 0.0%	100% 0.0%	0.0% 100%	90.7% 9.3%	98.7% 1.3%	96.7% 3.3%	0.0% 100%	100% 0.0%	86.7% 13.3%	100% 0.0%	0.0% 100%	100% 0.0%	91.3% 8.7%

**Classe Predita**

Figura 41. Matriz de confusão para o conjunto de teste, utilizando-se o método SGDM com 5 blocos de extração de características sem aumento de dados.

Ao analisarmos a matriz de confusão na Figura 41, é visto que a rede ainda apresenta dificuldades em aprender as características das classes: Batimento prematuro nodal (J), Batimento não classificado (Q), batida de fuga atrial (e) e Batimento prematuro nodal (j), analisando o resultado dessas classes é notado que não houve nenhum acerto no conjunto de testes.

Após o processo de treinamento utilizando o aumento de dados, o valor da acurácia global aumentou para 95,30% enquanto a acurácia média elevou-se para 89,10%. A Figura 42 mostra a nova curva de convergência da acurácia ao longo do treinamento, para os conjuntos de validação e teste.

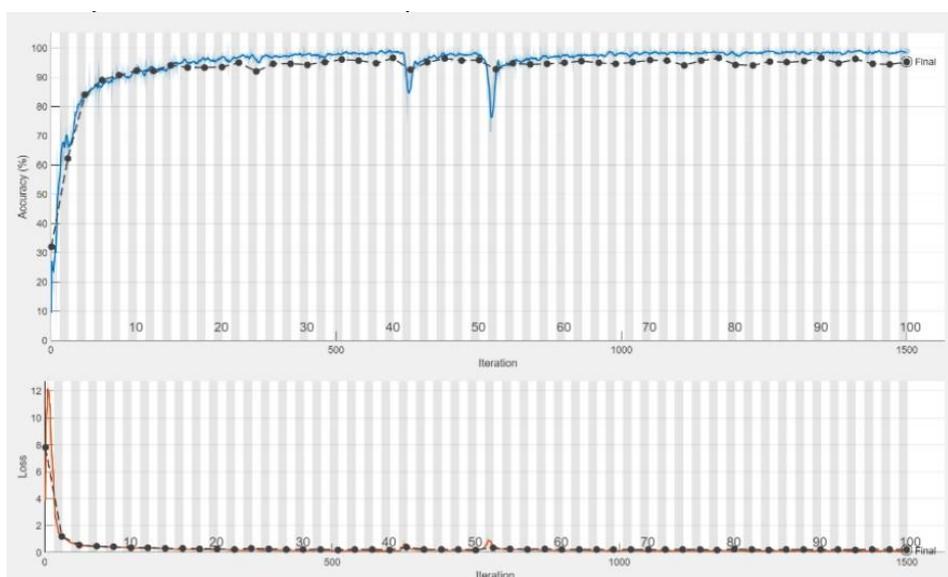


Figura 42. Curva de convergência da acurácia ao longo do treinamento, para os conjuntos de validação e teste, utilizando-se o método SGDM com 5 blocos de extração de características após o aumento de dados.

Ao analisarmos as curvas na Figura 42, são notados dois momentos em que ocorrem oscilações, próxima a 40ª e 50ª épocas, nota-se que a rede está estabilizada nas outras épocas, demonstrando que o processo de aprendizado está mais eficiente. A Figura 43 mostra a matriz de confusão obtida no conjunto de testes.

## Matriz de Confusão

Classe verdadeira	E	15 1.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%	
	F	0 0.0%	117 10.5%	0 0.0%	0 0.0%	0 0.0%	1 0.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	99.2% 0.8%
	J	0 0.0%	0 0.0%	13 1.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	L	0 0.0%	0 0.0%	0 0.0%	144 12.9%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	N	0 0.0%	0 0.0%	0 0.0%	0 0.0%	150 13.5%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	P	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	146 13.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	Q	0 0.0%	2 0.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	5 0.4%	0 0.0%	3 0.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	50.0% 50.0%
	R	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	148 13.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	V	0 0.0%	0 0.0%	0 0.0%	2 0.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	114 10.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	98.3% 1.7%
	a	0 0.0%	0 0.0%	0 0.0%	4 0.4%	0 0.0%	3 0.3%	0 0.0%	0 0.0%	33 3.0%	25 2.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	38.5% 61.5%
	e	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	3 0.3%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	f	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	2 0.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	147 13.2%	0 0.0%	98.7% 1.3%
	l	0 0.0%	1 0.1%	1 0.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	34 3.1%	94.4% 5.6%
			100% 0.0%	97.5% 2.5%	92.9% 7.1%	96.0% 4.0%	100% 0.0%	97.3% 2.7%	100% 0.0%	98.7% 1.3%	76.0% 24.0%	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%	95.3% 4.7%

## Classe Predita

Figura 43. Matriz de confusão para o conjunto de teste, utilizando-se o método SGDM com 5 blocos de extração de características após aumento de dados.

Ao analisarmos a nova matriz de confusão obtida após o aumento de dados na Figura 43, é visto que a rede superou as dificuldades em aprender as características das classes: Batimento prematuro nodal (J) agora com 92,9% de acerto, Batimento não classificado (Q) agora com 100% de acerto, batida de fuga atrial (e) e o Batimento prematuro nodal (j) agora ambos com 100% de acerto.

A terceira arquitetura, composta por 7 blocos de extração de características, observa-se que sem o aumento de dados o valor obtido da acurácia global foi 93,60% e a acurácia média foi de 53,83%. A Figura 44 apresenta a curva de convergência da acurácia ao longo do treinamento, para os conjuntos de validação e teste.

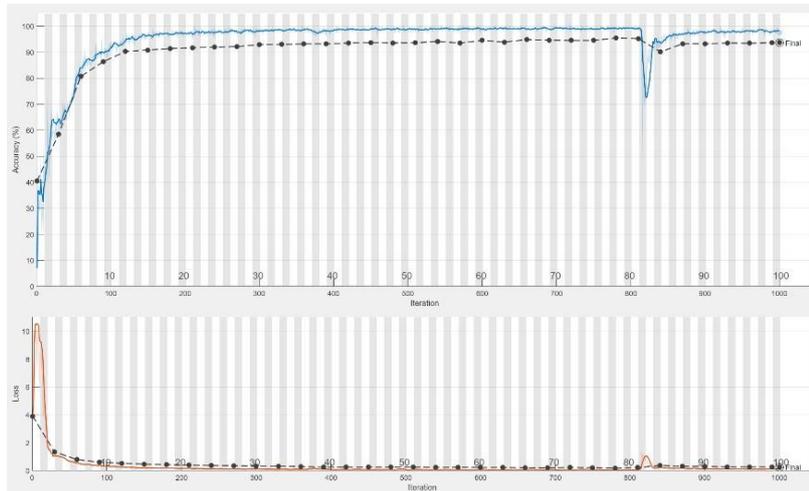


Figura 44. Curva de convergência da acurácia ao longo do treinamento, para os conjuntos de validação e teste, utilizando-se o método SGDM com 7 blocos de extração de características sem aumento de dados.

Observando a curva de convergência da acurácia na Figura 44, é possível notar que esta arquitetura com 7 blocos de extração de características já apresenta uma melhora significativa de desempenho mesmo sem o aumento de dados, ela se mantém estável a partir da 10ª época, apresenta uma oscilação próximo a 80ª época gerada a partir das classes com pouca quantidade de amostra. A seguir na Figura 45, é mostrada a matriz de confusão obtida no conjunto de testes.

### Matriz de Confusão

Classe verdadeira	E	15 1.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%	
	F	0 0.0%	118 10.6%	0 0.0%	3 0.3%	0 0.0%	3 0.3%	0 0.0%	0 0.0%	1 0.1%	1 0.1%	0 0.0%	0 0.0%	0 0.0%	93.7% 6.3%
	J	0 0.0%	0 0.0%	1 0.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	L	0 0.0%	0 0.0%	0 0.0%	146 13.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	N	0 0.0%	0 0.0%	4 0.4%	0 0.0%	150 13.5%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 0.1%	0 0.0%	87.7% 12.3%
	P	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	146 13.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	Q	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
	R	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 0.1%	150 13.5%	1 0.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	98.7% 1.3%
	V	0 0.0%	0 0.0%	8 0.7%	1 0.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	143 12.8%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	92.3% 7.7%
	a	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	24 2.2%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	e	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
	f	0 0.0%	0 0.0%	1 0.1%	0 0.0%	0 0.0%	1 0.1%	3 0.3%	0 0.0%	5 0.4%	0 0.0%	2 0.2%	147 13.2%	14 1.3%	85.0% 15.0%
	l	0 0.0%	2 0.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	2 0.2%	50.0% 50.0%
		100% 0.0%	98.3% 1.7%	7.1% 92.9%	97.3% 2.7%	100% 0.0%	97.3% 2.7%	0.0% 100%	100% 0.0%	95.3% 4.7%	96.0% 4.0%	0.0% 100%	100% 0.0%	5.9% 94.1%	93.6% 6.4%
	E	F	J	L	N	P	Q	R	V	a	e	f	l		
	Classe Predita														

Figura 45. Matriz de confusão para o conjunto de teste, utilizando-se o método SGDM com 7 blocos de extração de características sem aumento de dados.

Analisando a matriz de confusão na Figura 45, é possível notar que ainda existem 4 classes que a arquitetura não consegue extrair e aprender suas características são elas: Batimento prematuro nodal (J) 7,1% de acerto, Batimento não classificado (Q) 0 % de acerto, batida de fuga atrial (e) 0% de acerto e o Batimento prematuro nodal (j) com 5,9% de acerto enquanto as demais já possuem resultado satisfatório.

Após o processo de treinamento utilizando o aumento de dados, o valor da acurácia global aumentou para 95,80% enquanto a acurácia média elevou-se para 96,46%. A Figura 46 mostra a nova curva de convergência da acurácia ao longo do treinamento, para os conjuntos de validação e teste.

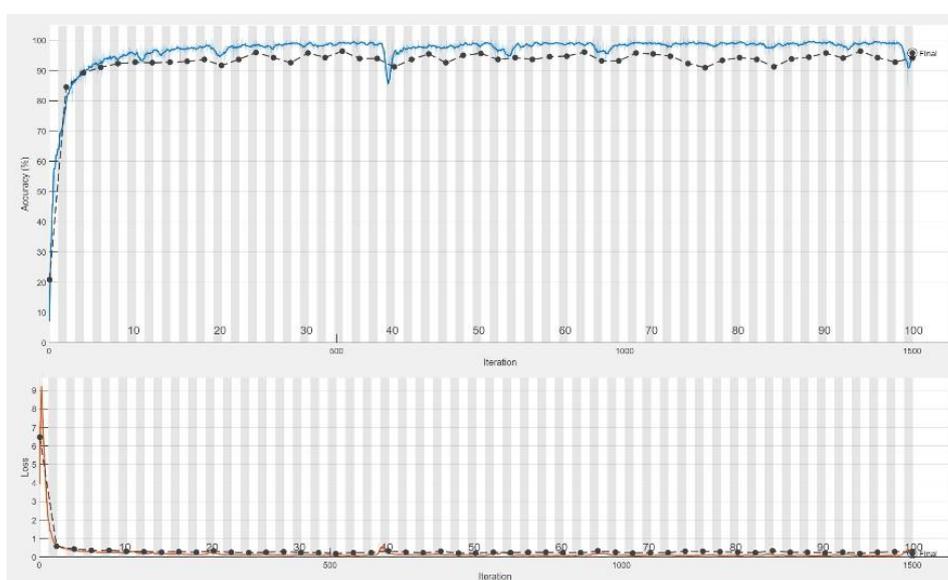


Figura 46. Curva de convergência da acurácia ao longo do treinamento, para os conjuntos de validação e teste, utilizando-se o método SGDM com 7 blocos de extração de características após aumento de dados.

Ao analisarmos as curvas na Figura 46, são notados dois momentos em que ocorrem oscilações, próxima a 40<sup>a</sup> e 100<sup>a</sup> épocas, nota-se que a rede está estabilizada nas outras épocas, demonstrando que o processo de aprendizado está mais eficiente. A Figura 47 apresenta a matriz de confusão obtida no conjunto de testes.

## Matriz de Confusão

Classe verdadeira	Classe Predita												Total			
	E	F	J	L	N	P	Q	R	V	a	e	f	L	Count	Accuracy	
E	15 1.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	15	100%	0.0%
F	0 0.0%	110 9.9%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	110	100%	0.0%
J	0 0.0%	0 0.0%	12 1.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 0.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	12	92.3%	7.7%
L	0 0.0%	0 0.0%	0 0.0%	148 13.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	148	100%	0.0%
N	0 0.0%	0 0.0%	0 0.0%	0 0.0%	150 13.5%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	150	100%	0.0%
P	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	146 13.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 0.1%	0 0.0%	146	99.3%	0.7%
Q	0 0.0%	2 0.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	5 0.4%	0 0.0%	7 0.6%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	12	35.7%	64.3%
R	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	150 13.5%	1 0.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	151	99.3%	0.7%
V	0 0.0%	0 0.0%	0 0.0%	2 0.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	122 11.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	124	98.4%	1.6%
a	0 0.0%	5 0.4%	1 0.1%	0 0.0%	0 0.0%	2 0.2%	0 0.0%	0 0.0%	18 1.6%	25 2.2%	0 0.0%	0 0.0%	0 0.0%	48	49.0%	51.0%
e	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	3 0.3%	0 0.0%	0 0.0%	3	100%	0.0%
f	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	2 0.2%	0 0.0%	0 0.0%	1 0.1%	0 0.0%	0 0.0%	146 13.1%	0 0.0%	149	98.0%	2.0%
L	0 0.0%	3 0.3%	1 0.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	34 3.1%	37	89.5%	10.5%
	100% 0.0%	91.7% 8.3%	85.7% 14.3%	98.7% 1.3%	100% 0.0%	97.3% 2.7%	100% 0.0%	100% 0.0%	81.3% 18.7%	100% 0.0%	100% 0.0%	99.3% 0.7%	100% 0.0%	95.8% 4.2%		

Figura 47. Matriz de confusão para o conjunto de teste, utilizando-se o método SGDM com 7 blocos de extração de características após aumento de dados.

Analisando a Figura 47, matriz de confusão gerada a partir do aumento de dados, é possível notar que as classes que anteriormente estavam com problema no processo de treinamento passaram a ter resultados satisfatórios, são elas: Batimento prematuro nodal (J) agora com 85,7% de acerto (antes com 7,1%), Batimento não classificado (Q) agora com 100% (antes com 0%), Batida de fuga atrial (e) agora com 100% de acerto (antes com 0%) e o Batimento prematuro nodal (j) agora com 100% de acerto (antes com 5,9%) as demais classes também mantiveram seus bons resultados.

Os resultados das três arquiteturas após a aplicação do aumento de dados foram reunidos e apresentados na Tabela 8.

Tabela 8. Resultados das arquiteturas propostas após o aumento de dados utilizando o método de otimização SGDM.

	<b>Arquiteturas Propostas</b>		
	<b>CNN2D -1</b>	<b>CNN2D -2</b>	<b>CNN2D -3</b>
<b>Acurácia Média</b>	78,96%	81,10%	<b>96,46%</b>
<b>Acurácia Global</b>	91,30%	95,30%	<b>95,80%</b>

A partir dos resultados apresentados na Tabela 8, concluímos que a CNN2D-3 é a arquitetura que melhor se adequou ao processo de aumento de dados com o otimizador SGDM. A partir disto, foi considerada essa arquitetura para obtenção das métricas de desempenho, as quais são apresentadas na Tabela 9.

Tabela 9. Desempenho do método proposto

<b>Arquitetura</b>	<b>Acurácia</b>	<b>Precisão</b>	<b>Sensibilidade</b>	<b>Especificidade</b>	<b>F1- Score</b>
CNN2D-3 com SGDM	95,80%	95,20%	91,67%	95,20%	93,40%

As métricas apresentadas na Tabela 9 podem ser interpretadas das seguintes maneiras:

- Acurácia: 95,80% das previsões da rede estão corretas;
- Precisão: A precisão de 95,20% indica que, quando a rede prevê uma classe, ela está correta na maioria das vezes. Isso sugere que as previsões positivas da rede são altamente confiáveis.
- Sensibilidade: A sensibilidade de 91,67% indica que a rede está capturando a maioria dos resultados positivos;
- Especificidade: A especificidade de 95,20% sugere que a rede está excluindo efetivamente resultados negativos;
- F1-Score: O F1-Score é uma métrica que leva em consideração tanto a precisão quanto a sensibilidade. O valor de 93,40% indica um bom equilíbrio entre essas duas métricas.

### 5.3 Resultados da Técnica de Visualização de ativação

O propósito da técnica de visualização de ativação é a identificação das regiões de um ECG onde a rede neural convolucional 2D (CNN2D) detectou uma arritmia cardíaca. Após o treinamento da CNN2D, aplicamos a técnica Grad-CAM a cada imagem no conjunto de testes. O Grad-CAM produziu mapas de ativação que destacam as áreas relevantes de cada imagem para cada classe específica. Em seguida, calculamos o coeficiente médio de ativação, o menor coeficiente e o desvio padrão para cada classe de arritmia. Esses cálculos são essenciais para compreender a importância média das regiões de cada imagem na previsão correta da classe.

Na Tabela 10 apresenta-se os valores de ativação de todos os pixels relevantes dos sinais/imagens de teste apresentados na entrada da rede, em função de sua classe, obtidos do mapa de ativação - Grad-CAM.

Tabela 10. Valores de Coeficientes médios para cada classe de arritmia no conjunto de testes.

	<b>Símbolo</b>	<b>Tipo de Evento</b>	<b>Coeficiente Médio por classe</b>	<b>Menor Coeficiente</b>	<b>Desvio padrão</b>
1	<b>E</b>	Batida de escape ventricular	0,980	0,910	0,028
2	<b>F</b>	Fusão de batimento ventricular e normal	0,870	0,800	0,034
3	<b>J</b>	Batimento prematuro nodal (juncional)	0,855	0,800	0,032
4	<b>L</b>	Batimento de bloqueio de ramo esquerdo	0,890	0,800	0,058
5	<b>P</b>	Ritmo compassado	0,880	0,800	0,059
6	<b>Q</b>	Batida inclassificável	0,920	0,860	0,031
7	<b>R</b>	Batida de bloqueio de ramo direito	0,920	0,880	0,022
8	<b>V</b>	Contração ventricular prematura	<b>0,842</b>	<b>0,802</b>	<b>0,033</b>
9	<b>a</b>	Batimento atrial prematuro aberrado	0,931	0,920	0,006
10	<b>e</b>	Batida de fuga atrial	0,923	0,892	0,015
11	<b>f</b>	Fusão de batimento compassado e normal	0,907	0,829	0,022
12	<b>j</b>	Batida de fuga nodal (juncional)	0,923	0,915	0,915

Ao analisar a Tabela 10 para determinar um valor mínimo de coeficiente do GradCAM que seja aplicável a todas as classes, consideramos várias abordagens. Uma delas foi a análise dos coeficientes médios de todas as classes para identificar um limiar global que capture as características gerais das regiões de interesse.

Após a análise, identificamos que o menor coeficiente médio tem um valor de 0,842, associado à classe de Contração Ventricular Prematura (V). Quanto ao menor valor de coeficiente, encontramos um valor de 0,8021, também atribuído à classe de Contração Ventricular Prematura (V). Com base nesse levantamento, adotamos o valor de 0,8 como um limiar mínimo para todas as classes. Assim, quanto em um dados pixel o valor no mapa de ativação for igual ou superior a 0,8 considera-se que a região ao qual esse pixel pertence foi relevante para a classificação. A seguir são apresentados os resultados obtidos a partir da aplicação dessa técnica. Na Figura 48, as faixas de cor verde marcam os trechos onde foram encontradas as arritmias. Nesse caso, trata-se a arritmia batida de escape ventricular (E).

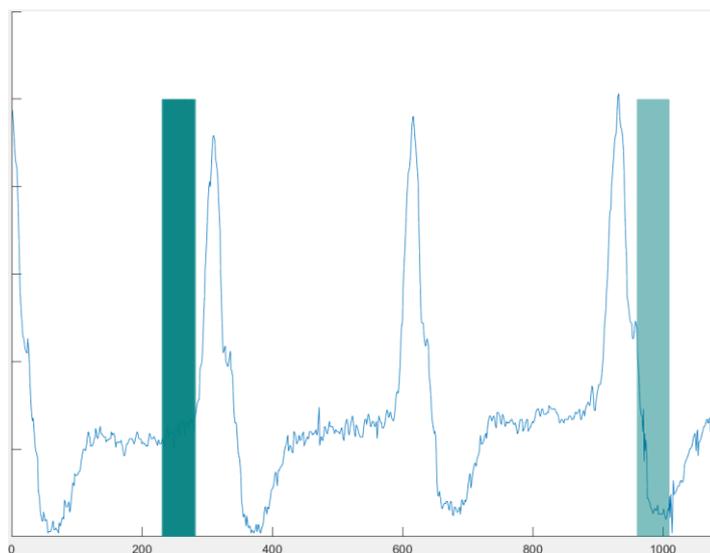


Figura 48. Subconjunto do sinal de ECG com destaque na corverde, mostrando a presença de dois pontos contendo a arritmia batida de escape ventricular (E).

Ao analisarmos a Figura 48, é possível notar que a rede identificou dois pontos contendo arritmia, ambos são marcados pela cor verde. A diferença de tonalidade das marcações representa o valor obtido através do mapa de pontuação, em termos práticos, quanto mais saturada for a cor maior foi o valor obtido no mapa.

A Figura 49 apresenta outro subconjunto de um sinal de ECG contendo três trechos com arritmia. Nesse caso, trata-se da arritmia fusão de batimento ventricular (F).

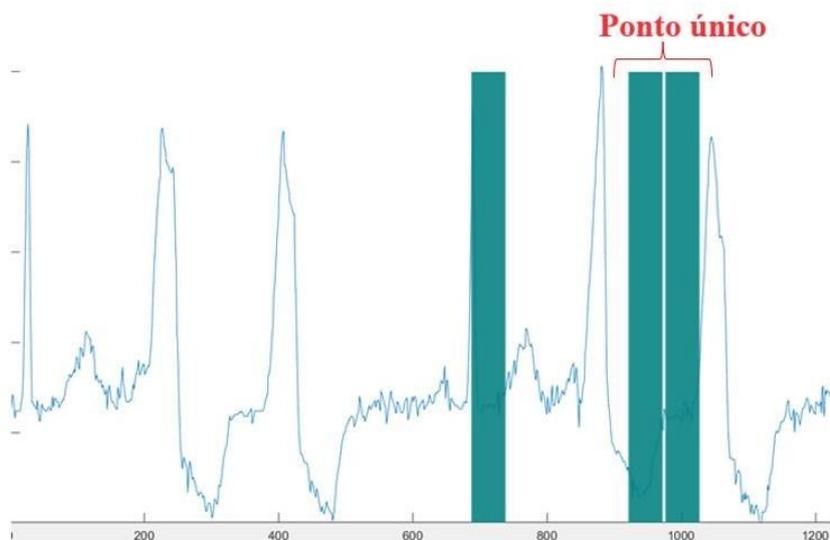


Figura 49. Subconjunto do sinal de ECG com destaque na cor verde, mostrando a presença da arritmia fusão de batimento ventricular (F).

Neste subconjunto, são mostrados 3 pontos com alta probabilidade de conter a arritmia fusão de batimento ventricular (F). Nota-se que os dois pontos localizados à direita, na verdade representam apenas uma arritmia.

A seguir nas Figuras 50, 51 e 52 temos, respectivamente, a identificação de trechos contendo as seguintes arritmias: bloqueio de ramo direito (R), contração ventricular prematura (V) e batida de fuga nodal (j).

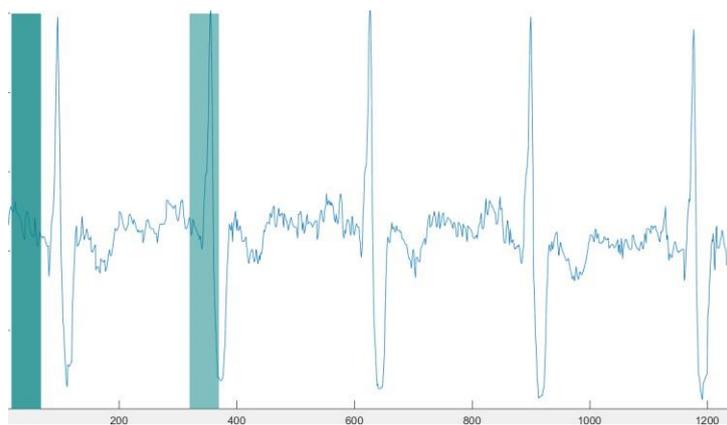


Figura 50 Subconjunto do sinal de ECG com destaque na cor verde, mostrando a presença de dois pontos contendo a arritmia bloqueio de ramo direito (R).

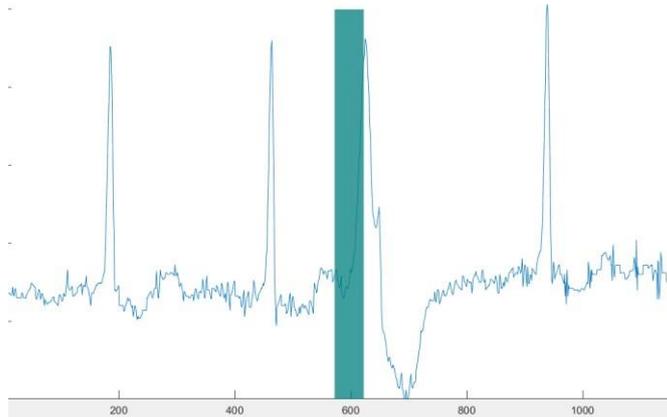


Figura 51. Subconjunto do sinal de ECG com destaque na cor verde, mostrando a presença da arritmia contração ventricular prematura (V).

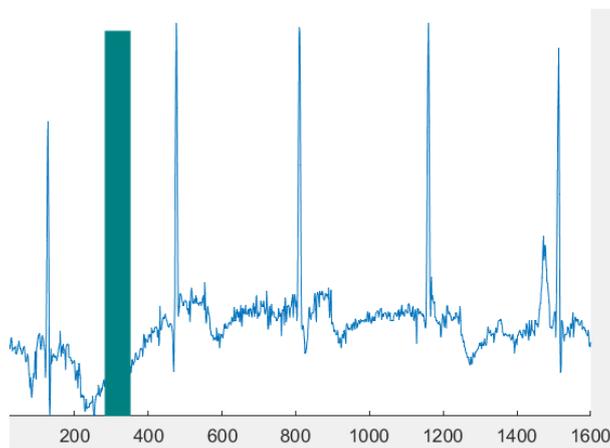


Figura 52. Subconjunto do sinal de ECG com destaque na cor verde, mostrando a presença da arritmia batida de fuga nodal (j).

## 6. CONCLUSÃO

Neste estudo, foram introduzidas três variações da arquitetura de rede neural Convolucional 2D, denominadas CNN2D-1, CNN2D-2 e CNN2D-3. A distinção entre elas reside na composição dos blocos de extração de características *Conv-Batch-ReLu*, sendo respectivamente 3, 5 e 7 blocos de extração. Observou-se que a arquitetura CNN2D-3 alcançou o melhor desempenho, destacando-se devido à sua maior quantidade de blocos de extração de características. Durante as simulações, evidenciou-se que o acréscimo de camadas convolutivas e etapas de subamostragem resultou na geração de mais mapas de características, permitindo que a rede aprendesse de maneira mais eficaz a diferenciar entre as classes.

No contexto das simulações, foram empregados dois métodos de otimização: ADAM e SGDM. Ambos foram associados a ajustes de hiper parâmetros. O otimizador SGDM apresentou resultados superiores, manifestados por curvas mais suaves durante as etapas de treinamento e teste. Em contrapartida, o método ADAM não se adaptou eficientemente à arquitetura proposta, resultando em acurácias notadamente baixas ou até nulas em ambos os conjuntos de treinamento e teste.

Quanto à técnica de aumento de dados proposta, destacou-se o aumento de desempenho em quatro classes durante os testes na melhor arquitetura CNN2D-3. Sem o aumento de dados, a acurácia para Batimento prematuro nodal (J) era de 7,1%, Batimento não classificado (Q) 0%, Batida de fuga atrial (e) 0%, e Batimento prematuro nodal (j) 5,9%. Após a aplicação do aumento de dados, esses resultados aumentaram significativamente para 85,7%, 100%, 100% e 100%, respectivamente. Isso valida a eficácia da técnica de aumento de dados, cujo ponto forte é a preservação das características originais do sinal de ECG.

Com a técnica de visualização de ativação aplicada, o GradCAM, verificou-se que os mapas de pontuação gerados destacaram visualmente os trechos contendo as arritmias identificadas e classificadas pela rede, quando as pontuações do referido mapa eram iguais ou superiores a 0,8. Essa visualização é relevante por proporcionar uma informação visual e não apenas a classificação da rede.

Finalmente, ressalta-se a contribuição deste estudo no uso de redes convolucionais 2D para identificação e classificação de arritmias cardíacas em sinais de ECG. A geração de imagens de intensidade a partir da conversão dos valores da amostra de ECG para tons

de cinza, juntamente com a aplicação de uma nova técnica de aumento de dados, demonstram avanços. Ao comparar com trabalhos publicados, nosso estudo superou o estado da arte ao obter uma acurácia de 95,80% na melhor arquitetura CNN2D-3, em comparação com 95,20% de YILDIRIM *et al.* com 13 classes semelhantes. Acreditamos que a utilização de sinais bidimensionais, permitindo a exploração de uma vizinhança maior, e o método de aumento de dados proposto, foram responsáveis pela melhoria nos resultados. A avaliação das três arquiteturas propostas e dos métodos de otimização contribuiu para a determinação da melhor arquitetura.

Para trabalhos futuros, sugere-se a exploração e aplicação de Redes Neurais Recorrentes de Memória de Longo Prazo (LSTM) e suas variações, visando aprimorar ainda mais os resultados obtidos.

## 7. REFERÊNCIAS

ACHARYA, U. R., FUJITA, H., OH, et al. **Deep convolutional neural network for the automated diagnosis of congestive heart failure using ECG signals.** Information Sciences, 415-416, 190-198. 2017.

ACHARYA, U. R., OH et al. **Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals.** Computers in Biology and Medicine, 100, 270-278. 2018.

AMERICAN HEART ASSOCIATION. **"Life's Simple 7."** Disponível em: <https://www.heart.org/en/healthy-living/healthy-lifestyle/my-life-check--lifes-simple-7>

AMERICAN HEART ASSOCIATION. **Arrhythmia.** Disponível em: [https://www.heart.org/en/health-topics/arrhythmia\\_\(2021\)](https://www.heart.org/en/health-topics/arrhythmia_(2021))

AMERICAN HEART ASSOCIATION. **Arrhythmia.** Disponível em: <https://www.heart.org/en/health-topics/arrhythmia>. Acesso em: 08/05/2023.

Abdalla, H. I., & Amer, A. A. (2022). **Towards Highly-Efficient k-Nearest Neighbor Algorithm for Big Data Classification.** In Proceedings of the 5th International Conference on Networks, Information Systems, and Security: Predict Intelligent Systems in Interconnected Digital Worlds based on 5G/6G (NISS), Bandung, Indonesia (pp. 1-5). doi: 10.1109/NISS55057.2022.10085013.

BEYER, S. E., MOLL, J. N., & PUMAROLA, M. **ECG: An Introductory Course: A Practical Introduction to Clinical Electrocardiography.** Springer.2019.

BOSCARIOLI, C., DA SILVA, I. N., & PERES, et al. **Uma abordagem histórica dos sistemas neurais artificiais e sua aplicação na previsão do comportamento do consumo de energia elétrica.** Sistemas & Gestão, 11(3), 285-297. 2016.

COSTA, L. F., & LUDWIG JÚNIOR, A. **Redes neurais artificiais: fundamentos e aplicações**. LTC Editora S.A. 2007.

C. M. Ugwu, C. Pierrette Mukamakuza and E. Tuyishimire, "**ECG-Signals-based Heartbeat Classification: A Comparative Study of Artificial Neural Network and Support Vector Machine Classifiers**," 2024 IEEE 22nd World Symposium on Applied Machine Intelligence and Informatics (SAMI), Stará Lesná, Slovakia, 2024, pp. 000217-000222, doi: 10.1109/SAMI60510.2024.10432834.

DATA SCIENCE ACADEMY. **Deep Learning Book**. Disponível em: <<https://www.deeplearningbook.com.br/>>. Acesso em: 10/05/2023.

FRANÇA, M. S. **Doença cerebrovascular**. In **Tratado de neurologia da Academia Brasileira de Neurologia**. Guanabara Koogan. (2018).

GUPTA, A., et al. **Performance Evaluation of Support Vector Machines and Artificial Neural Networks for the Classification of Cardiac Arrhythmias**. Journal of Medical Systems, 43(9), 1-10. 2019.

G. Kumar, S. K. Pandey, N. Varshney, R. R. Janghel, K. U. Singh e A. Kumar, "**Arrhythmia Detection from ECG Signals using CNN Model**" apresentado na 6ª Conferência Internacional sobre Sistemas de Informação e Redes de Computadores (ISCON), Mathura, Índia, 2023, pp. 1-6, doi: 10.1109/ISCON57294.2023.10112173.

HELDER JOSÉ LIMA REIS. et al. **ECG: manual prático de eletrocardiograma**. editora-- São Paulo : Editora Atheneu, 2013.

JONES, M. **Deep learning architectures: The rise of artificial intelligence**.

Disponívelem:<https://developer.ibm.com/articles/cc-machine-learning-deep-learning-architectures/>.2017.

J. Nandanwar, J. Singh and S. Patidar, "ECG Signals- Early detection of Arrhythmia using Machine Learning approaches," 2023 13th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 2023, pp. 32-38, doi: 10.1109/Confluence56041.2023.10048810.

LECUN, Y., BENGIO, Y., & HINTON, G. **Deep learning**. Nature, 521(7553), 436-444. 2015.

MACFARLANE, P. W., et al. **Variability in ECG interpretation in routine clinical practice**. Journal of Electrocardiology, 51(6), 997-1002. 2018,

MACHADO, P. G., COSTA, I. M. et al. **Detecção precoce de doenças cardiovasculares utilizando técnicas computacionais**. In Anais do XXIII Simpósio Brasileiro de Informática na Educação (SBIE) (pp. 124-133). 2019.

MINISTÉRIO DA SAÚDE. **Datasus - mortalidade**. Disponível em: <<http://www2.datasus.gov.br/DATASUS/index.php?area=0205&id=6942>> .2020.

NATIONAL HEART, LUNG, AND BLOOD INSTITUTE. **"What Is the Heart?"** Disponível em: <<https://www.nhlbi.nih.gov/health-topics/what-is-the-heart>> 2018.

NATIONAL INSTITUTES OF HEALTH. **"Heart Anatomy."** Disponível em: <<https://medlineplus.gov/ency/article/002305.htm>>

RAJKOMAR, A., et al. **Scalable and Accurate Deep Learning with Electronic Health Records**. npj Digital Medicine, 1(1), 1-10. 2018.

RAJPURKAR, et al. **Cardiologist-level arrhythmia detection with convolutional neural networks**. arXiv preprint arXiv:1707.01836. 2017.

RAJPURKAR, P., et al. **Cardiologist-Level Arrhythmia Detection and Classification in Ambulatory Electrocardiograms Using a Deep Neural Network**. Nature Medicine, 25(1), 65-69. 2017.

SHAH, R. S., et al. **Artificial Intelligence in Arrhythmias and Electrocardiography**. Arrhythmia & Electrophysiology Review, 9(4), 251-259. 2020.

SIMOES, M. V., OLIVEIRA, J. L., & MELO, P. L. **Cardiac arrhythmia classification using 2D and 1D convolutional neural networks**. In Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC) (Vol. 5, pp. 430-437). 2018.

SWEIS, R., & JIVAN, A. **Coronary artery disease.** In StatPearls [Internet]. StatPearls Publishing. 2020.

THOMAS, D., & KHETRAPAL, S. **Rheumatic Fever and Rheumatic Heart Disease.** In StatPearls [Internet]. StatPearls Publishing.

UNIVERSITY OF MARYLAND MEDICAL CENTER. "**Heart Anatomy.**" Disponível em: <<https://www.umms.org/ummc/healthservices/heartascular/services/cardiovascular-health/heart-anatomy>>.

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

PAN, J.; TOMPKINS, W. J. **A real-time QRS detection algorithm.** IEEE Transactions on Biomedical Engineering, v. BME-32, n. 3, p. 230-236, mar. 1985.