

Classificação de Doença Arterial Coronariana através de redes neurais profundas

Diogo Casal Ferreira¹, Alexei Manso Correa Machado^{1,2}

¹Departamento de Ciência da Computação

Pontifícia Universidade Católica de Minas Gerais (PUC Minas), Belo Horizonte, Brasil

²Departamento de Anatomia e Imagem - Faculdade de Medicina

Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Brasil

diogocasalf@gmail.com, alexeimcmachado@gmail.com

Abstract. *Coronary artery disease (CAD) is one of the leading causes of global mortality, highlighting the need for more effective early diagnosis methods. This study compares LSTM, Bi-GRU, and Transformer neural networks in classifying electrocardiograms (ECG) into three categories: normal, inferior myocardial infarction (IMI), and anterior myocardial infarction (AMI), using the PTB-XL dataset. Results show that Transformers offer faster training (average of 35 minutes), while LSTM and Bi-GRU achieve higher accuracy (91% and 92%, respectively). The study also contributes to improving ECG data preprocessing and training for cardiac diagnosis classification.*

Resumo. *As doenças arteriais coronarianas (DAC) figuram entre as principais causas de mortalidade no mundo, demandando avanços no diagnóstico precoce. Este estudo compara redes neurais LSTM, Bi-GRU e Transformer na classificação de eletrocardiogramas (ECG) em três categorias: normal, infarto inferior (IMI) e anterior (AMI), utilizando o conjunto de dados PTB-XL. Os resultados indicam que Transformers são mais rápidos no treinamento (média de 35 minutos), enquanto LSTM e Bi-GRU apresentam maior precisão (91% e 92%, respectivamente). O trabalho também contribui para a melhoria do pré-processamento e do treinamento de dados de ECG aplicados ao diagnóstico cardíaco.*

1. Introdução

O eletrocardiograma (ECG) é uma ferramenta essencial na avaliação de condições cardíacas por sua acessibilidade e custo relativamente baixo em comparação com outros exames. Ele desempenha um papel importante na detecção precoce, monitoramento e classificação de doenças cardiovasculares, como a doença arterial coronariana (DAC), uma das principais causas de mortalidade e morbidade no mundo. Segundo [Komilovich (2023)], a DAC é responsável por 32,7% das doenças cardiovasculares, resultando em um custo anual superior a 200 bilhões de dólares para os sistemas de saúde dos Estados Unidos. A detecção precoce é fundamental para prevenir complicações, e o ECG, por ser amplamente disponível, emerge como um recurso valioso para esse propósito, permitindo intervenções antecipadas que podem salvar vidas.

O uso de redes neurais (RNs) revolucionou o diagnóstico assistido em diversas áreas médicas, especialmente na análise de dados sequenciais como os de ECG. Modelos como LSTMs (Long Short-Term Memory), Bi-GRU (Bi-Directional Gated Recurrent

Unit) e Transformers demonstraram alta eficácia ao processar grandes volumes de dados clínicos, identificando padrões complexos que muitas vezes escapam à análise humana.

Poucos trabalhos na literatura tratam da comparação direta entre técnicas e redes de processamento de sequência mais antigas, como LSTM e Bi-GRU, com abordagens mais atuais, como Transformers, especialmente no contexto da análise de sinais biomédicos. Portanto, o objetivo desta pesquisa é comparar a eficiência de diferentes modelos de redes neurais – LSTM, Bi-GRU e Transformer – que utilizam diferentes estratégias para a classificação de doença arterial coronariana, utilizando o conjunto de dados PTB-XL desenvolvido por [Wagner et al. (2020)]. Além disso, busca-se propor um modelo que se destaque em termos de desempenho, contribuindo para diagnósticos mais acessíveis, precisos e confiáveis. Essa análise não apenas avalia as capacidades individuais dos modelos, mas também explora como o uso de inteligência artificial pode transformar o diagnóstico médico e melhorar os desfechos para pacientes com DAC.

2. Trabalhos Relacionados

[Darmawahyuni et al. (2019)] utilizaram LSTMs na classificação de infarto do miocárdio (MI) a partir de sinais de ECG, com a aplicação da técnica de segmentação de sinais em janelas de 4 segundos, com pelo menos quatro batimentos cardíacos por janela, para que fossem usados como treinamento para a rede neural. O modelo foi treinado com diversas camadas, variando de uma a três, todas otimizadas com o algoritmo Adam. Ainda em 2019, [Zhang et al. (2019b)] propuseram um trabalho que se destacou pelas técnicas de pré-processamento. Entre elas, o uso do algoritmo Pan-Tompkins, proposto por [Pan e Tompkins (1985)], que permitiu a extração de batimentos cardíacos a partir dos dados brutos de ECG, identificando o complexo QRS de cada batimento. Ainda utilizando LSTM, [Dey et al. (2021)] inovaram e estruturaram um modelo de rede neural simples que combinava uma camada de CNN 1-D e uma camada LSTM bidirecional, uma variante da LSTM que processa os dados sequenciais em duas direções, para tratar sinais de ECG.

No campo das GRUs, [Zhang et al. (2019a)] apresentam um estudo que explorou o uso de uma GRU bidirecional para a captura de uma maior quantidade de contexto temporal em comparação ao GRU simples. Ademais, utilizou o algoritmo Pan-Tompkins para filtragem e ampliação dos dados do dataset. Recentemente, por meio de uma análise sistemática da literatura, [Geng et al. (2023)] ampliaram essa abordagem ao proporem uma rede neural multitarefa que combinava Bi-GRU com módulos SE-ResNet e Transformer. No pré-processamento, a frequência dos sinais de ECG foi reduzida de 500 Hz para 250 Hz, acelerando o treinamento. Os dados pre-processados encaminhados para o módulo Bi-GRU, que extraí características das sequências de ECG em duas direções. A função de perda final combina as perdas das tarefas principal e auxiliar, ajustadas de acordo com o modelo multitarefa.

Mais recentemente, o uso da arquitetura de Transformers introduziu novas perspectivas na análise de ECGs, como no estudo proposto por [Natarajan et al. (2020)], que utilizou uma rede de transformadores ampla e profunda para classificar ECGs multidi-derivacionais, demonstrando desempenho satisfatório em termos de velocidade de iteração e treinamento do modelo em comparação a outros métodos. Contudo, não apresentaram resultados melhores que as outras RNNs, comprovando que a eficiência de sua aplicação depende de diversos fatores. Um aprimoramento significativo foi proposto por [Atiea e

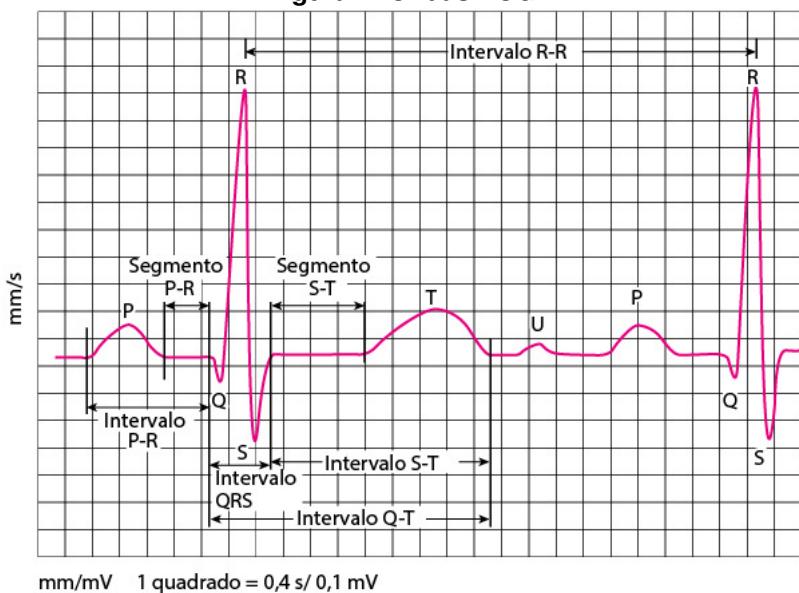
Adel (2022)], que trouxe uma melhora no desempenho ao combinar características morfológicas e dados brutos de ECG. Outro trabalho relevante foi produzido por [Butt et al. (2022)], que explorou o uso de um modelo híbrido CNN-LSTM combinado com o mecanismo de Attention/Transformer para a análise de ECGs de 12 derivações. Esse modelo híbrido conseguiu melhorar o resultado dos modelos de atenção, ao combinar as forças de CNN, LSTM, superando as limitações de modelos convencionais, evitando o *overfitting* em conjuntos de dados médios e mostrando-se eficaz na classificação precisa de ECGs.

Por fim, [Xiong et al. (2022)] apresentam uma análise de *Deep Learning* (DL) para detecção e localização de infarto do miocárdio (MI), e apresentaram estudos sobre a arquitetura LSTM e GRU, trazendo insights significativos que guiaram o estudo na análise de identificação de doenças arteriais coronárias de um modo geral. Contudo, se limitaram a redes neurais recorrentes, e não extenderam a análise à tecnologias atuais, como os módulos de atenção dos Transformers.

3. O Eletrocardiograma

O eletrocardiograma (ECG) é uma ferramenta diagnóstica essencial para monitorar a atividade elétrica do coração. Ele registra essa atividade por meio de eletrodos colocados na pele, gerando um gráfico de dados sequenciais que mostra as diferentes ondas que representam cada fase do ciclo cardíaco. As principais ondas analisadas no ECG (Figura 1) são: a onda P, o complexo QRS, o segmento ST, a onda T e, em alguns casos específicos, a onda U. Essas ondas fornecem informações essenciais sobre a saúde do coração e são comumente utilizadas para detectar arritmias, distúrbios na condução elétrica e, especialmente, eventos isquêmicos, como o infarto do miocárdio, que pode se manifestar por elevações ou depressões no segmento ST, inversões da onda T e presença de ondas Q patológicas. Esses segmentos são mostrados na Figura 1.

Figura 1. Ondas ECG



Fonte: <https://www.msdmanuals.com/pt/casa/multimedia/figure/ecg-como-interpretar-as-ondas>

4. Redes neurais para séries temporais

As Redes Neurais Recorrentes (RNNs) são arquiteturas projetadas especificamente para lidar com dados sequenciais, como séries temporais, textos, sinais de áudio e sinais biomédicos, como os eletrocardiogramas. Diferentemente das redes neurais tradicionais, as RNNs têm a capacidade de manter uma memória dos estados anteriores, o que é essencial quando a ordem ou dependência temporal dos dados é importante. As RNNs clássicas, no entanto, enfrentam desafios, como o problema do *vanishing gradient*, onde as contribuições de eventos longínquos na sequência se tornam insignificantes durante o treinamento. Para superar essas limitações, variantes foram desenvolvidas, como as redes LSTM (Long Short-Term Memory) e GRU (Gated Recurrent Unit). Além disso, arquiteturas como os Transformers se destacaram por sua capacidade de lidar com grandes sequências de forma mais eficiente.

4.1. Long Short-Term Memory (LSTM)

As Redes LSTM foram introduzidas por [Hochreiter (1997)] para mitigar as limitações das RNNs tradicionais, especialmente em tarefas que exigiam a captura de dependências de longo prazo. A LSTM incorpora um sistema de portas (*gate units*) que controla o fluxo de informações dentro da rede. Essas portas incluem a porta de entrada, a porta de esquecimento e a porta de saída, que permitem à rede adicionar ou remover informações de sua memória, garantindo que apenas informações relevantes sejam retidas ao longo do tempo. Essa estrutura torna a LSTM eficaz em tarefas que envolvem longas dependências temporais, como a análise de sinais de ECG, onde eventos anômalos podem ser espaçados por várias amostras de dados.

4.2. Bidirectional Gated Recurrent Unit (Bi-GRU)

A Rede Bi-GRU, publicada por [Chung (2014)], é uma variação das RNNs que combinam a simplicidade das GRUs com a bidirecionalidade, ou seja, a capacidade de processar uma sequência em ambas as direções, capturando não apenas o contexto passado, mas também o futuro em relação a cada ponto da sequência. As GRUs, como as LSTMs, utilizam portas para controlar o fluxo de informações, mas possuem uma estrutura mais simples e leve, o que pode levar a um treinamento mais rápido. No caso da Bi-GRU, a vantagem reside na habilidade de considerar tanto o histórico, quanto o futuro de um determinado ponto da sequência, tornando-a útil para análises como a predição de doenças a partir de ECGs.

4.3. Transformers

Proposto por [Vaswani (2017)], o Transformer é uma arquitetura de rede neural que introduz o conceito de *self-attention*. Diferente das RNNs, que processam sequências de forma linear, os Transformers podem considerar todas as posições da sequência ao mesmo tempo, permitindo uma análise muito mais eficiente e paralelizada de grandes quantidades de dados. Em vez de depender de uma estrutura sequencial, os Transformers utilizam o mecanismo de atenção para identificar automaticamente quais partes da sequência são mais relevantes para a tarefa de predição. Isso os torna especialmente adequados para lidar com dependências de longo alcance em grandes conjuntos de dados. No contexto da análise de eletrocardiogramas, essa arquitetura permite uma maior flexibilidade e precisão na detecção de padrões dos sinais elétricos.

5. Metodologia

5.1. Base de Dados

O primeiro conjunto de dados utilizado neste trabalho é o PTB-XL, que consiste em 21.799 registros clínicos de eletrocardiogramas (ECG) de 12 derivações, provenientes de 18.869 pacientes. Cada registro, com 10 segundos de duração, foi coletado sob a supervisão do Physikalisch-Technische Bundesanstalt (PTB)[Wagner et al. (2020)]. Além de informações demográficas, os registros incluem 12 derivações e detalhes como eixo cardíaco e estágio do infarto.

O segundo conjunto de dados utilizado é o dataset PTB-XL+ [Strothoff et al. (2023)], um dataset complementar ao PTB-XL, possuindo uma coleção abrangente de dados de ECG complementados por metadados adicionais, como batimentos medianos e pontos fiduciais, além de declarações diagnósticas. Os dados oferecem uma visão abrangente da atividade elétrica do coração, com 71 estados diferentes de ECG. Identificadores únicos para ECG e pacientes facilitam a organização e análise.

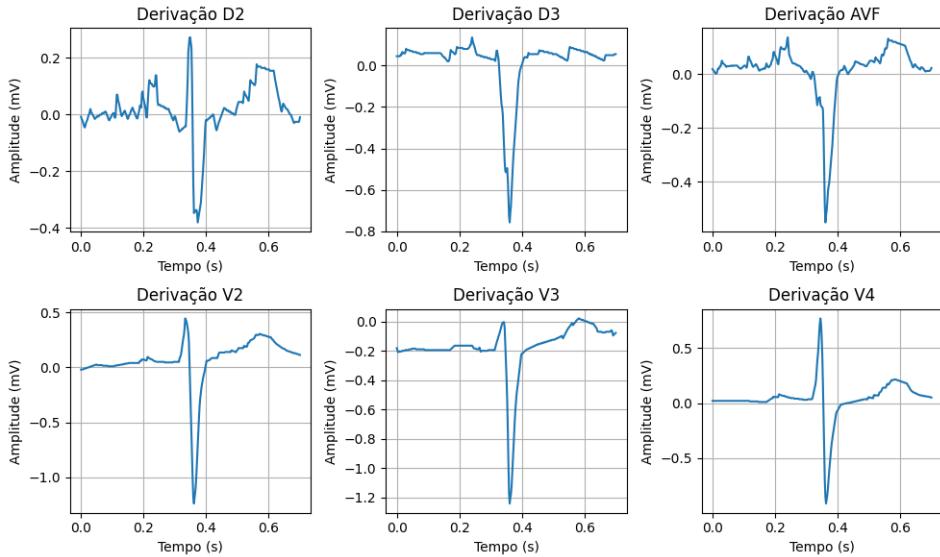
5.2. Pré-processamento

O pré-processamento dos dados foi realizado utilizando a linguagem Python no ambiente Jupyter Notebook, com o auxílio da biblioteca Pandas. Foram selecionadas as classes relacionadas a doença arterial coronariana – infarto do miocárdio anterior (AMI), infarto do miocárdio inferior (IMI) e dados normais (NORM) – totalizando 7.427 registros de 4.655 pacientes distintos. Registros duplicados de pacientes foram removidos do conjunto de testes para evitar redundâncias, mas mantidos no conjunto de treinamento para aumentar a quantidade de dados disponíveis.

Foram selecionadas apenas as derivações mais relevantes para a classificação dos tipos de infarto. As derivações D2, D3 e aVF, obtidas por eletrodos posicionados nos membros, representam a atividade elétrica no plano frontal inferior e são essenciais para detectar infarto do miocárdio inferior. Já as derivações V2, V3 e V4, registradas por eletrodos colocados no tórax, refletem a atividade no plano horizontal anterior, sendo fundamentais para identificar infarto do miocárdio anterior.

O algoritmo Pan-Tompkins de [Pan e Tompkins (1985)], foi usado para segmentar os sinais brutos, identificando o complexo QRS nos registros de ECG através da técnica de janelamento, que consiste em dividir o sinal em segmentos menores, ou janelas, cada uma contendo apenas um batimento cardíaco. Essa abordagem permitiu a extração de cinco batimentos isolados por registro, sendo cada um deles, utilizado como entrada para a rede separadamente, ampliando a quantidade de dados e concedendo um maior enfoque em padrões morfológicos das ondas. Após a segmentação, foi aplicado o ajuste de curva polinomial, uma técnica que utiliza funções polinomiais para modelar e remover o desvio de linha de base presente nos sinais, garantindo maior precisão nas análises subsequentes. Além disso, foi utilizado um filtro de rejeição de banda (*notch filter*), projetado para eliminar interferências em frequências específicas, como ruídos provenientes da rede elétrica, melhorando ainda mais a qualidade dos dados extraídos.

Figura 2. Exemplo da extração do algoritmo Pan-Tompkins



5.3. Treinamento

As redes neurais empregadas neste estudo foram desenvolvidas e treinadas a partir do estado inicial, sem a utilização de pesos previamente ajustados, abrangendo três arquiteturas distintas: LSTM, Bi-GRU e Transformer. Os treinamentos foram realizados em um ambiente com processador Intel Core i5-13450H (10 núcleos), 16 GB de RAM DDR5 e GPU NVIDIA GeForce RTX 3050 com 6 GB de memória dedicada, sob o sistema operacional Windows 11. Em todas as redes, a camada de saída utilizou a função de ativação **Softmax**, amplamente utilizada para problemas de classificação multiclasse. Essa função converte os valores da saída em probabilidades, garantindo que a soma das probabilidades para todas as classes seja igual a 1:

$$\text{Softmax}(z_i) = \frac{e^{z_i}}{\sum_{j=1}^n e^{z_j}}$$

onde z_i é o valor de ativação da i -ésima classe, e n é o número total de classes.

A função de perda utilizada foi a **categorical cross-entropy**, ideal para problemas de classificação multiclasse. Essa função mede a diferença entre a distribuição prevista pelo modelo e a verdadeira distribuição dos dados.

$$\text{Loss} = - \sum_{i=1}^n y_i \log(\hat{y}_i)$$

onde y_i é o rótulo verdadeiro (1 para a classe correta e 0 para as demais), e \hat{y}_i é a probabilidade prevista para a i -ésima classe.

O algoritmo Adam foi escolhido como otimizador devido à sua eficiência e adaptabilidade durante o treinamento. Para avaliar o desempenho das redes ao longo das épocas, foi utilizada a métrica de acurácia, que mede a proporção de previsões corretas em relação ao total de amostras avaliadas. Outro mecanismo utilizado para otimizar o processo de

treinamento e evitar problemas de sobreajuste, foi uma função personalizada que utilizava *early stopping* com uma paciência de 15 épocas e também realizava o salvamento automático do modelo treinado, da matriz de confusão e dos gráficos de treinamento, incluindo as informações mais relevantes sobre a arquitetura utilizada. Esses recursos facilitaram uma análise mais aprofundada dos resultados e auxiliaram no ajuste fino dos modelos.

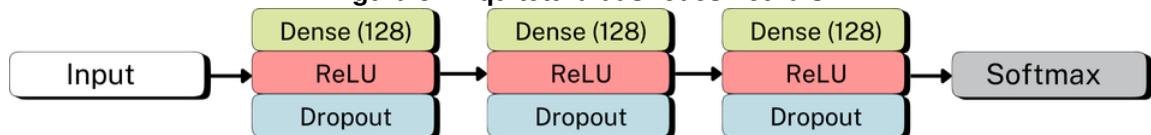
5.4. Arquitetura

A rede foi desenvolvida sem o uso de pesos pré-treinados, recebendo como entrada séries temporais de ECG tratadas pelo algoritmo de extração Pan-Tompkins e emitindo previsões de suas categorias. Durante o processo de ajuste de hiperparâmetros, diferentes tamanhos de camadas ocultas foram testados, variando entre 32, 64, 128 e 256 neurônios. A configuração com 32 neurônios foi insuficiente para o aprendizado do modelo, enquanto 256 neurônios resultaram em *overfitting*, indicando que a rede estava se ajustando demais aos dados de treinamento e não lidando bem com dados desconhecidos. Entre as configurações de 64 e 128 neurônios, a última se destacou, oferecendo um equilíbrio superior entre desempenho e generalização, tornando-se a escolha ideal para a arquitetura final.

Pela elevada quantidade de neurônios em sua estrutura, foi possível extrair mais características relevantes durante o treinamento, como por exemplo as alterações no segmento ST, essencial para a detecção de infarto, como discutido por [Carneiro (1997)]. Cada camada densa oculta é intercalada com uma camada de *dropout*, visando melhorar a generalização do modelo, reduzir o *overfitting* e aumentar sua robustez na predição.

A arquitetura detalhada se encontra na figura 3 e se resume a 3 camadas densas ocultas com 128 neurônios cada + *dropout* após cada camada + camada softmax com 3 saídas. A função de ativação utilizada nas camadas ocultas foi a ReLU (Rectified Linear Unit), que acelera a convergência e ajuda a mitigar problemas de gradiente. A camada softmax na saída fornece as probabilidades das classes, permitindo a classificação final dos sinais de ECG.

Figura 3. Arquitetura das redes neurais



5.5. Seleção de Modelos

A comparação com os modelos LSTM, Bi-GRU e Transformer para dados sequenciais foi motivada pela natureza temporal e complexa dos sinais de ECG, que exigem a captura de padrões ao longo do tempo para uma classificação precisa. A LSTM e o Bi-GRU são arquiteturas de Redes Neurais Recorrentes (RNNs) que possuem a capacidade de lidar com dependências de longo prazo, sendo muito utilizadas em tarefas envolvendo dados sequenciais. Já o Transformer, com sua arquitetura baseada em mecanismos de atenção, destaca-se como uma abordagem atual e inovadora por processar sequências em paralelo, ao invés de forma sequencial como as redes recorrentes. Por conta disso, se tornou comumente utilizada em diversos problemas de aprendizado de máquina.

5.5.1. Comparação de Modelos

Para a comparação dos modelos, foram considerados critérios como acurácia, precisão, *recall* e F1-score, para uma avaliação mais abrangente dos modelos, e eficiência computacional por sua importância em cenários clínicos de decisões rápidas e uso otimizado de recursos, além da demonstração da matriz de confusão para cada caso. Cada modelo foi treinado e avaliado no mesmo conjunto de dados de ECG, com a análise de seu desempenho sendo feita em função da capacidade de categorização das diferentes variações de sinais cardíacos. Além disso, para cada modelo, foi gerada a matriz de confusão, que apresenta as classificações corretas e incorretas em relação às categorias definidas, permitindo uma análise mais detalhada do desempenho em cada classe.

6. Resultados Experimentais

Nesta seção, é realizada uma análise comparativa do desempenho dos modelos LSTM, Bi-GRU e Transformer sob as mesmas circunstâncias. O uso do algoritmo Pan-Tompkins possibilitou aumentar a quantidade de dados, tendo em vista que a fornecida pelo *dataset* com seus dados primários não foi suficiente para a classificação das DACs. Nos experimentos realizados para classificação de eletrocardiogramas em três classes (NORMAL, IMI e AMI), as redes neurais Transformer, LSTM e Bi-GRU apresentaram um desempenho não muito variado nas métricas de sensibilidade, F1-score, precisão e especificidade, mas com leves alterações e diferenças notáveis de tempo de treinamento. Cada uma dessas métricas foi calculada considerando o desempenho de cada rede em identificar corretamente as classes sob as mesmas condições de treinamento: número de épocas igual a 100, taxa de aprendizado de 0,0001, 3 camadas densas com 128 neurônios e um *dropout* de 0,5. Observando as tabelas de métricas, a rede Bi-GRU destacou-se em sensibilidade, mostrando maior capacidade em identificar corretamente as classes IMI e AMI, o que é crítico em contextos médicos. Já no F1-score e precisão, as redes Bi-GRU e LSTM, similarmente, tiveram um desempenho mais equilibrado entre classes, evidenciando sua robustez em situações de classes desbalanceadas. Todas essas informações podem ser verificadas nas Figuras 4, 5 e 6.

Figura 4. Tabela com os resultados do treinamento da rede Bi-GRU

CLASSES	SENSIBILIDADE	ESPECIFICIDADE	PRECISÃO	SCORE F1
AMI	0,93	0,98	0,95	0,94
IMI	0,92	0,95	0,90	0,91
NORM	0,91	0,95	0,91	0,91

Figura 5. Tabela com os resultados do treinamento da rede LSTM

CLASSES	SENSIBILIDADE	ESPECIFICIDADE	PRECISÃO	SCORE F1
AMI	0,92	0,97	0,94	0,93
IMI	0,89	0,96	0,92	0,90
NORM	0,93	0,93	0,89	0,91

Figura 6. Tabela com os resultados do treinamento do Transformer

CLASSES	SENSIBILIDADE	ESPECIFICIDADE	PRECISÃO	SCORE F1
AMI	0,87	0,97	0,92	0,90
IMI	0,83	0,94	0,86	0,85
NORM	0,91	0,90	0,84	0,88

O histórico do treinamento, ilustrado pelos gráficos de acurácia e perda da Figura 7, aponta que a rede Transformer convergiu de forma mais rápida, iniciando o treinamento desde a primeira época com 88% de acurácia no treino e de 86% no teste, sugerindo uma adaptação mais eficiente ao conjunto de dados utilizado, além de ter obtido uma média de 35 minutos em seu tempo de treinamento, enquanto as redes LSTM e Bi-GRU obtiveram 50 minutos e 1 hora respectivamente. A linha vermelha presente na Figura 7 indica o desempenho final do modelo de inteligência artificial nos dados de teste. Esses dados consistem em registros de ECGs que não foram utilizados durante o treinamento da rede neural, sendo completamente desconhecidos por ela. Pode-se observar que as redes Bi-GRU e LSTM se destacam em sua capacidade de manter uma alta precisão em generalizar as classificações para novos dados, refletindo em uma aplicação prática mais segura em cenários reais.

As matrizes de confusão, visualizadas nas Tabelas 1, 2, 3, por outro lado, indicam que a rede LSTM apresentou menores taxas de falsos positivos e falsos negativos para a classe normal, enquanto a Bi-GRU teve desempenho superior nas classes de infarto. Os resultados obtidos refletem as características esperadas de cada arquitetura. O Transformer, conhecido por se destacar em cenários com grandes volumes de dados, não apresentou vantagem significativa neste caso, devido à quantidade moderada de dados do conjunto PTB-XL, favorecendo arquiteturas como LSTM e Bi-GRU.

Tabela 1. Matriz de Confusão para método LSTM

	AMI	IMI	Normal
AMI	613	26	25
IMI	14	612	63
Normal	24	29	721

Tabela 2. Matriz de Confusão para método Bi-GRU

	AMI	IMI	Normal
AMI	619	15	30
IMI	16	634	39
Normal	18	55	701

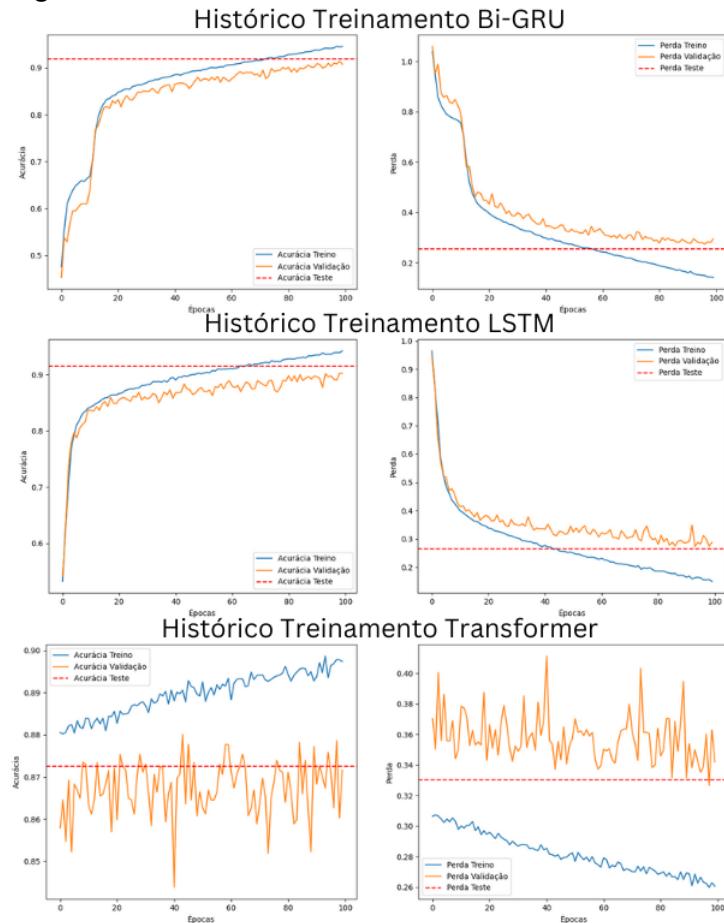
Tabela 3. Matriz de Confusão para método Transformer

	AMI	IMI	Normal
AMI	577	40	47
IMI	30	573	86
Normal	17	51	706

7. Conclusão

A comparação entre Bi-GRU, LSTM, Transformer na classificação de eletrocardiogramas revelou que cada arquitetura possui suas próprias vantagens em contextos específicos. Em geral, a rede Bi-GRU demonstrou um bom desempenho em sensibilidade e especificidade, tornando-a uma ótima escolha em aplicações que buscam a precisão no diagnóstico de condições cardíacas. Por outro lado, o LSTM obteve resultados moderados em várias métricas, sendo uma opção de escolha em cenários onde é necessário um desempenho estável entre diferentes classes. Já o Transformer destacou-se por seu menor tempo de treinamento e rápida convergência, o que é vantajoso em aplicações que exigem rapidez e escalabilidade, lidando muito bem com grandes volumes de dados.

Figura 7. Tabela com os resultados do treinamento



Dado o contexto clínico, a escolha da arquitetura ideal depende das prioridades de aplicação - se o objetivo for minimizar falsos negativos em diagnósticos de alto risco, a rede Bi-GRU pode ser preferida; para uma abordagem mais equilibrada, a LSTM também se mostra promissora. As condições de treinamento e a análise com métricas variadas permitiram uma comparação objetiva, destacando como as redes neurais respondem a diferentes desafios em dados de ECG.

Para trabalhos futuros, sugere-se a exploração de arquiteturas híbridas que combinem os pontos fortes de LSTM, Bi-GRU e Transformer que foram levantados neste estudo, buscando uma maior eficácia na análise de sinais de ECG. Adicionalmente, seria relevante aplicar técnicas de atenção explicável(XAI) para melhorar a interpretabilidade dos resultados e facilitar a aplicação clínica e funcional das redes neurais propostas. Outro aspecto promissor é ampliar o escopo de estudo para contemplar mais subclasse da DAC, como infarto do miocárdio lateral, posterior, apical, septal, subendocárdico e transmural, possibilitando uma melhor compreensão das especificidades de cada caso e um diagnóstico mais preciso e abrangente para condições cardíacas complexas.

Finalmente, seria de grande utilidade a integração dessa tecnologia de classificação diretamente aos softwares utilizados nos eletrocardiogramas. Com essa implementação, durante o exame, o médico poderia ter acesso imediato a *insights* relevantes, como a identificação preliminar de padrões associados a diferentes tipos de infarto ou outras anormalidades cardíacas.

Agradecimentos — AMC Machado agradece o auxílio financeiro do Fundo de Incentivo à Pesquisa FIP-PUCMinas 2025/32467 e da FAPEMIG através dos projetos APQ-02753-24 e APQ-06556-24.

Referências

- M. A. Atiea e M. Adel. Transformer-based neural network for electrocardiogram classification. *International Journal of Advanced Computer Science and Applications*, 13 (11), 2022.
- F. S. Butt, M. F. Wagner, J. Schäfer, e D. G. Ullate. Toward automated feature extraction for deep learning classification of electrocardiogram signals. *IEEE Access*, 10:118601–118616, 2022.
- E. F. Carneiro. O eletrocardiograma: 10 anos depois. In *O Eletrocardiograma: 10 anos depois*, pages 622–622. 1997.
- J. Chung. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*, 2014.
- A. Darmawahyuni, S. Nurmaini, e Sukemi. Deep learning with long short-term memory for enhancement myocardial infarction classification. In *2019 6th International Conference on Instrumentation, Control, and Automation (ICA)*, pages 19–23, 2019. doi: 10.1109/ICA.2019.8916683.
- M. Dey, N. Omar, e M. A. Ullah. Temporal feature-based classification into myocardial infarction and other cvds merging cnn and bi-lstm from ecg signal. *IEEE Sensors Journal*, 21(19):21688–21695, 2021. doi: 10.1109/JSEN.2021.3079241.
- Q. Geng, H. Liu, T. Gao, R. Liu, C. Chen, Q. Zhu, e M. Shu. An ecg classification method based on multi-task learning and cot attention mechanism. *Healthcare*, 11(7), 2023. ISSN 2227-9032. doi: 10.3390/healthcare11071000. URL <https://www.mdpi.com/2227-9032/11/7/1000>.

- S. Hochreiter. Long short-term memory. *Neural Computation MIT-Press*, 1997.
- E. B. Komilovich. Coronary artery disease. *EUROPEAN JOURNAL OF MODERN MEDICINE AND PRACTICE*, 3(12):81–87, 2023.
- A. Natarajan, Y. Chang, S. Mariani, A. Rahman, G. Boverman, S. Vij, e J. Rubin. A wide and deep transformer neural network for 12-lead ecg classification. In *2020 Computing in Cardiology*, pages 1–4. IEEE, 2020.
- J. Pan e W. J. Tompkins. A real-time qrs detection algorithm. *IEEE Transactions on Biomedical Engineering*, BME-32(3):230–236, 1985. doi: 10.1109/TBME.1985.325532.
- N. Strothoff, T. Mehari, C. Nagel, P. J. Aston, A. Sundar, C. Graff, J. K. Kanters, W. Haverkamp, O. Dössel, A. Loewe, et al. PtB-XL+, a comprehensive electrocardiographic feature dataset. *Scientific data*, 10(1):279, 2023.
- A. Vaswani. Attention is all you need. *arXiv preprint arXiv:1706.03762*, 2017.
- P. Wagner, N. Strothoff, R.-D. Bousseljot, D. Kreiseler, F. I. Lunze, W. Samek, e T. Schaeffer. PtB-XL, a large publicly available electrocardiography dataset. *Scientific data*, 7 (1):1–15, 2020.
- P. Xiong, S. M.-Y. Lee, e G. Chan. Deep learning for detecting and locating myocardial infarction by electrocardiogram: A literature review. *Frontiers in cardiovascular medicine*, 9:860032, 2022.
- X. Zhang, R. Li, H. Dai, Y. Liu, B. Zhou, e Z. Wang. Localization of myocardial infarction with multi-lead bidirectional gated recurrent unit neural network. *IEEE Access*, 7: 161152–161166, 2019a. doi: 10.1109/ACCESS.2019.2946932.
- X. Zhang, R. Li, Q. Hu, B. Zhou, e Z. Wang. A new automatic approach to distinguish myocardial infarction based on lstm. In *2019 8th International Symposium on Next Generation Electronics (ISNE)*, pages 1–3, 2019b. doi: 10.1109/ISNE.2019.8896550.