

doi <http://dx.doi.org/10.18265/2447-9187a2022id7597>  
ARTIGO ORIGINAL

SUBMETIDO 14/03/2023

APROVADO 17/04/2023


PUBLICADO ON-LINE 07/05/2023


PUBLICADO 10/10/2024

EDITOR ASSOCIADO  
Pedro Bertemes Filho

# Sim-ECG: plataforma computacional de código aberto para classificação de patologias cardíacas

 Helder Alves Pereira <sup>[1]</sup> \*

 João Pedro dos Santos Silva <sup>[2]</sup>

 Pedro Henrique dos Santos Almeida <sup>[3]</sup>

 Letícia Chaves Lima Cananéa <sup>[4]</sup>

[1] [helder.pereira@dee.ufcg.edu.br](mailto:helder.pereira@dee.ufcg.edu.br)

[2] [joao.pds.silva@ee.ufcg.edu.br](mailto:joao.pds.silva@ee.ufcg.edu.br)

[3] [pedro.almeida@ee.ufcg.edu.br](mailto:pedro.almeida@ee.ufcg.edu.br)

[4] [leticia.cananea@ee.ufcg.edu.br](mailto:leticia.cananea@ee.ufcg.edu.br)

Centro de Engenharia Elétrica e  
Informática, Universidade Federal  
de Campina Grande (UFCG),  
Campina Grande, Paraíba, Brasil

\* Autor para correspondência.

**RESUMO:** Técnicas de inteligência artificial vêm sendo utilizadas no auxílio ao diagnóstico de patologias. Com elas, torna-se possível analisar automaticamente uma grande quantidade de dados complexos. Dessa forma, patologias podem ser detectadas em estágios menos avançados. Neste artigo, apresenta-se uma plataforma computacional de código aberto que analisa sinais obtidos por meio de exames de eletrocardiograma. Para isso, foram utilizadas bases de dados públicas que possibilitaram analisar e classificar as informações, considerando diversos métodos de inteligência artificial. Os resultados obtidos demonstram a viabilidade da plataforma para estudos e desenvolvimento de ferramentas computacionais para auxílio no diagnóstico precoce de patologias cardíacas.

**Palavras-chave:** aprendizagem de máquina; aprendizagem profunda; classificação; eletrocardiograma; inteligência artificial.

## *Sim-ECG: An open-source software for classification of cardiac pathologies*

**ABSTRACT:** Artificial intelligence techniques have been used to aid in the diagnosis of pathologies. With them, it becomes possible to automatically analyze a large number of complex data. Pathologies can be detected in less advanced stages. This paper presents an open-source computational platform that analyzes signals obtained through electrocardiogram exams. For this, public databases have been used that made possible to analyze and classify the information, considering various methods of artificial intelligence. The results obtained demonstrate the feasibility of the platform for studies and development of computational tools to aid in the early diagnosis of cardiac pathologies.

**Keywords:** artificial intelligence; classification; deep learning; electrocardiogram; machine learning.



## 1 Introdução

Segundo a Sociedade Brasileira de Arritmias Cardíacas (SOBRAC), a arritmia cardíaca é “a alteração que ocorre na geração ou na condução do estímulo elétrico do coração e pode provocar modificações do ritmo cardíaco” (SOBRAC, 2022). A falta de identificação e tratamento sistemáticos dessa condição crônica pode levar a infarto, doença cardíaca e morte súbita (Fletcher, 2021).

Uma das formas de realizar o diagnóstico de arritmias cardíacas é por meio do exame de eletrocardiograma (ECG). Esse exame registra oscilações elétricas resultantes dos batimentos do coração, a partir de eletrodos distribuídos pelo corpo do paciente (Ferreira; Lopes, 2013).

É importante verificar que a análise do estado cardíaco, por meio do exame de ECG, não é uma tarefa trivial para o processo de identificação de doenças. De fato, o sinal de ECG não é estacionário e, portanto, os sintomas de uma doença, se houver, podem não ocorrer regularmente (Ferreira; Lopes, 2013). Para minimizar os problemas ocasionados pela análise manual de sinais de ECG, muitos trabalhos na literatura abordam o uso de técnicas de inteligência artificial para detectar as anomalias presentes nesse tipo de sinal (Assen *et al.*, 2020; Feeny *et al.*, 2020; Jiang *et al.*, 2020; Kachuee; Fazeli; Sarrafzadeh, 2018; Kilic, 2020; Siontis *et al.*, 2021).

Nesse contexto, apresenta-se, neste artigo, o desenvolvimento do Sim-ECG<sup>1</sup>, um simulador de código aberto (*software open-source*) para detecção de anomalias em conjuntos de dados disponíveis em bases de dados públicas, utilizando exames de ECG de derivação única. Com o Sim-ECG, é possível utilizar diferentes bases de dados, considerando diversos métodos de inteligência artificial, obtendo-se resultados satisfatórios em comparação com resultados disponíveis na literatura (Dey; Chaudhuri; Munshi, 2018; Guvenir *et al.*, 1997; Hsieh *et al.*, 2020; Li *et al.*, 2018; Urtnasan *et al.*, 2018). As bases de dados utilizadas neste artigo consideram a formatação dos arquivos do tipo .csv, em que as primeiras colunas representam as características do sinal e a última coluna é, em formato numérico, a classe a que pertence. Essas bases de dados estão disponíveis no repositório do Sim-ECG<sup>1</sup> e podem ser utilizadas pelos usuários para simulação dos cenários descritos na seção 4.

Este artigo está organizado da seguinte forma: na seção 2, apresenta-se uma breve descrição do estado da arte sobre métodos de inteligência artificial aplicados na classificação de sinais cardíacos. Na seção 3, apresentam-se o Sim-ECG e suas principais características. Na seção 4, descrevem-se os cenários utilizados para validação do simulador. Na seção 5, apresenta-se uma análise sobre os resultados obtidos para cada um dos cenários avaliados e, por fim, na seção 6, as conclusões.

## 2 Revisão da literatura

Li, Rajagopalan e Clifford (2013) utilizaram três bases de dados diferentes para fazer diagnósticos de fibrilação ventricular e taquicardia por meio do eletrocardiograma. O primeiro passo da abordagem foi o desenvolvimento, no qual foi extraída do ECG uma janela específica de 5 segundos com a utilização de um algoritmo genético para gerar a combinação ideal de variáveis para treino e teste com o método de máquina de vetor de suporte (SVM – *Support Vector Machine*). As combinações que resultaram no melhor desempenho foram selecionadas para a próxima fase, em que foram utilizadas diferentes durações de janela (de 1 segundo a 10 segundos). A eficácia obtida foi de 98,1%.

[1] Disponível em: <https://github.com/helderufcg/Sim-ECG/releases/tag/v1.0.0>. Acesso em: 7 maio 2023.

Rahman *et al.* (2015) tiveram como objetivo identificar, automaticamente, pacientes com cardiomiopatia hipertrófica (HC – *Hypertrophic Cardiomyopathy*). Foram utilizados sinais padrão com janela de 10 segundos e 12 derivações de ECG de pacientes para treinar e testar classificadores de batimentos cardíacos, tais como floresta aleatória (RF – *Random Forest*) e SVM. Como base de dados, foram considerados dois grupos de pacientes cardiovasculares: i) 221 pacientes com HC (com um total de 754 sinais de ECG); ii) 541 indivíduos, todos diagnosticados com cardiomiopatia isquêmica, ou não isquêmica, e com desfibrilador cardioversor implantável (ICD – *Implantable Cardioverter Defibrillator*), usados como controle. Como resultado, foram obtidas: i) precisão de aproximadamente 85%; ii) sensibilidade de 90%; iii) especificidade de 90%. Também foi comparado o desempenho desses dois classificadores (RF e SVM) com o obtido pelo classificador de regressão logística (LRC – *Logistic Regression Classifier*), porém os dois primeiros métodos tiveram um desempenho melhor.

Li *et al.* (2018) propuseram o uso dos métodos de aprendizagem de máquina (ML – *Machine Learning*), como SVM, RF e  $k$  vizinhos mais próximos ( $k$ NN – *k Nearest Neighbors*), em comparação com a utilização do modelo de cadeia de Markov, para diagnóstico de taquicardia supraventricular. Concluíram que, apesar de o modelo de cadeia de Markov ser facilmente aplicado em dados coletados por meio de ECG, esse apresentou uma eficácia 18,1% menor em relação aos métodos de ML, que obtiveram 91,1% de eficácia.

Dey, Chaudhuri e Munshi (2018) propuseram um método de detecção da apneia obstrutiva do sono (AOS), que se baseava em uma estrutura de aprendizagem profunda (DL – *Deep Learning*), empregando uma rede neural convolucional (CNN – *Convolutional Neural Network*) e utilizando sinais de ECG, disponíveis em uma base de dados pública (Moody; Mark, 2000). Uma parte da rede proposta realizava aprendizagem de recursos e a outra parte classificava os recursos de maneira supervisionada. Esse algoritmo obteve acurácia de 98,91%, sensibilidade de 97,82% e especificidade de 99,20%.

Em outro estudo relevante, Hsieh *et al.* (2020) apresentaram um método de detecção automática de fibrilação atrial (AF – *Atrial Fibrillation*), baseado em uma arquitetura de CNN de uma dimensão. O objetivo foi melhorar a precisão da detecção, reduzir a complexidade da rede e o custo computacional em relação a redes já existentes. Para o processo de treinamento, foi utilizado o conjunto de dados fornecidos pela *PhysioNet Challenge 2017*, que contém gravações de ECG com diferentes durações. A arquitetura apresentava dez blocos convolucionais, duas camadas totalmente conectadas e uma camada *softmax* como previsão de saída. De acordo com os resultados obtidos, o método atingiu uma pontuação média de todas as classes (F1) de 78,2%.

Na literatura, foram propostos diversos algoritmos que utilizam inteligência artificial para classificação de sinais de ECG, considerando diferentes patologias (Dey; Chaudhuri; Munshi, 2018; Guvenir *et al.*, 1997; Hsieh *et al.*, 2020; Li; Rajagopalan; Clifford, 2013; Li *et al.*, 2018; Rahman *et al.*, 2015; Urtnasan *et al.*, 2018) e abordagens diversas (Assen *et al.*, 2020; Feeny *et al.*, 2020; Jiang *et al.*, 2020; Kilic, 2020; Siontis *et al.*, 2021). Entretanto, é possível observar que nenhuma das referências trata do desenvolvimento de um *software open-source* para classificação de patologias, que apresente interface gráfica para o usuário. No Quadro 1, apresenta-se um resumo com as contribuições da literatura, em termos de uso de bases de dados públicas, aplicações, técnicas de inteligência artificial utilizadas e métricas de desempenho.

**Quadro 1 ►**

Contribuições da literatura, em termos de uso de bases de dados públicas, aplicações, técnicas de inteligência artificial utilizadas e métricas de desempenho.

Fonte: elaborado pelos autores

Referência	Base de dados	Aplicação	Técnica	Métrica de desempenho
Li; Rajagopalan; Clifford (2013)	<i>American Heart Association</i>  <i>Creighton University Ventricular Tachyarrhythmia</i>  <i>MIT-BIH Malignant Ventricular Arrhythmia</i>	Fibrilação ventricular e taquicardia	Algoritmo genético e SVM	Eficácia
Rahman <i>et al.</i> (2015)	–	Cardiomiopatia hipertrófica	LRC, RF e SVM	Precisão, sensibilidade e especificidade
Li <i>et al.</i> (2018)	–	Taquicardia ventricular	kNN, RF e SVM	Eficácia
Dey, Chaudhuri e Munshi (2018)	<i>Apnea-ECG</i>	Apneia obstrutiva do sono	CNN	Acurácia, sensibilidade e especificidade
Hsieh <i>et al.</i> (2020)	<i>PhysioNet Challenge 2017</i>	Fibrilação atrial	CNN	F1

### 3 Sim-ECG

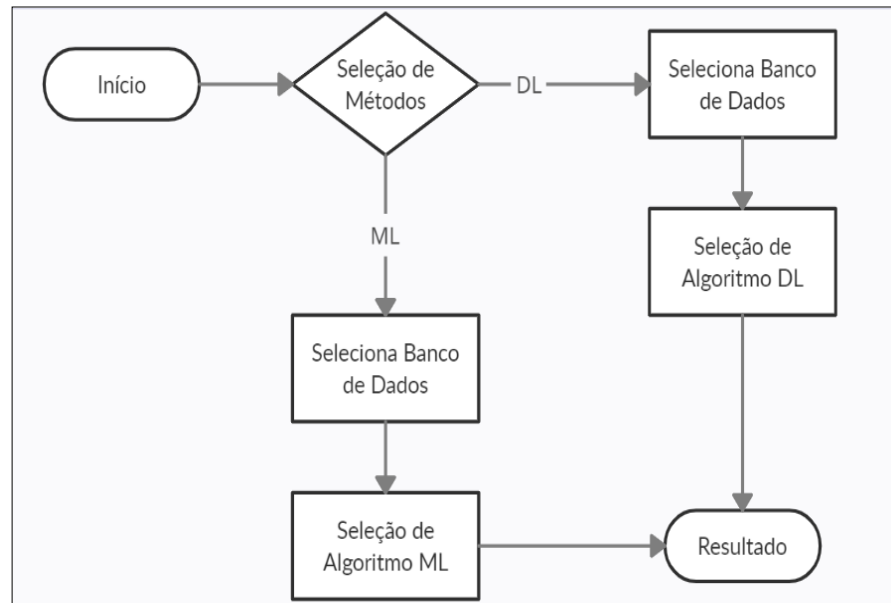
O Sim-ECG é capaz de utilizar diferentes métodos de classificação, entre eles SVM, kNN, *Naive Bayes* (NB), intervalos de características de votação (VFI – *Voting Feature Intervals*) e CNNs, por exemplo. Além de utilizar diferentes métodos de classificação, são consideradas diferentes bases de dados já pré-processadas e disponíveis na literatura, tais como: PTB Diagnostic ECG (Bousseljot, 2004), Apnea-ECG (Moody; Mark, 2000), MIT-BIH Arrhythmia (Moody; Mark, 2005). O processo de classificação consiste em três etapas (Dey; Chaudhuri; Munshi, 2018; Guvenir *et al.*, 1997; Hsieh *et al.*, 2020; Li; Rajagopalan; Clifford, 2013; Li *et al.*, 2018; Rahman *et al.*, 2015; Urtnasan *et al.*, 2018); são elas: i) pré-processamento do sinal de ECG; ii) extração de características; iii) classificação dos batimentos cardíacos. Com os dados pré-processados e disponibilizados na literatura, é possível desenvolver um software que não demanda uma quantidade considerável de espaço de memória, diminuindo sua complexidade e tornando-o didático com relação à aplicação de conceitos de inteligência artificial.

O Sim-ECG foi desenvolvido na linguagem de programação Python e utiliza os seguintes pacotes: *Matplotlib* (visualização de gráficos), *Scikit-learn* (aplicação de algoritmos de ML), *Pandas* (manipulação de banco de dados), *Numpy* (tratamento específico de dados), *TensorFlow* (aplicação de algoritmos de DL) e *Tkinter* (desenvolvimento de interface gráfica).

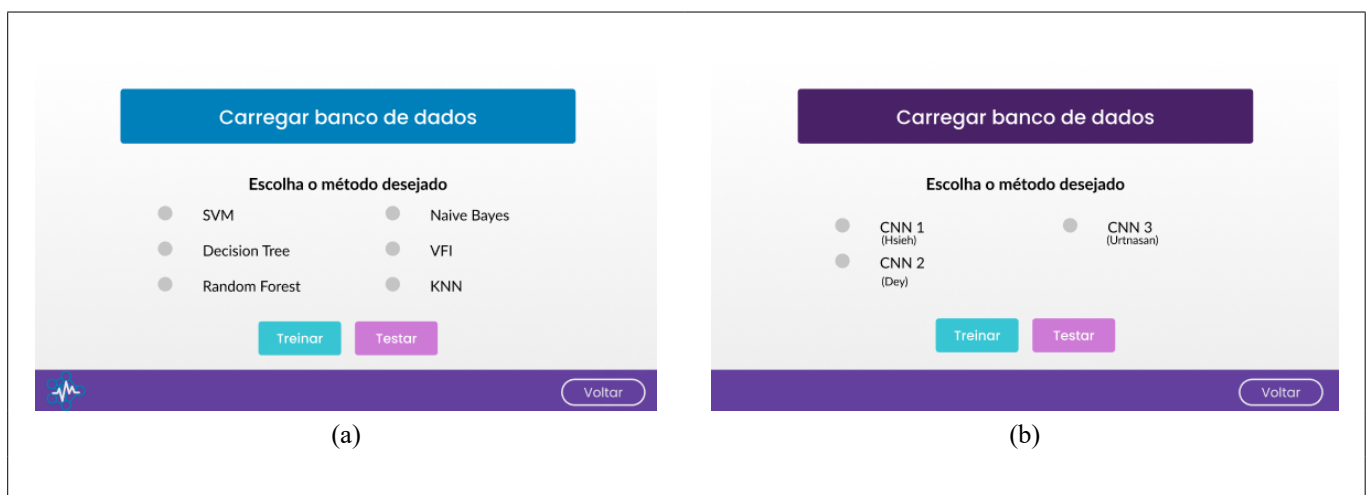
O fluxograma do Sim-ECG está ilustrado na Figura 1. Na tela de início, o usuário pode escolher ver informações sobre o simulador, o que o levará para a tela de informações, ou iniciar o processo de treinamento de um determinado método, o que o levará para a tela com a seleção dos métodos de ML (Figura 2a) ou DL (Figura 2b). Após a seleção do método desejado, o usuário deve carregar uma base de dados ou banco de dados. Em seguida, inicia-se o treinamento do modelo com o classificador selecionado. Ao final dessa etapa, serão exibidos os resultados, na tela de resultados. Por fim, o usuário pode voltar para a tela inicial e repetir todo o processo. Com o Sim-ECG, é possível obter a

acurácia de cada dobra – separações independentes do mesmo conjunto de dados –, bem como a média aritmética da eficácia, considerando todas as  $k$  dobras, conforme ilustrado na Figura 3. Também é possível obter resultados em termos de erros do modelo, ou perdas, e acurácias, ou assertividade, de treinamento em função do número de épocas, conforme ilustrado na Figura 4.

**Figura 1 ▶**  
Fluxograma do Sim-ECG.  
Fonte: elaborado pelos autores

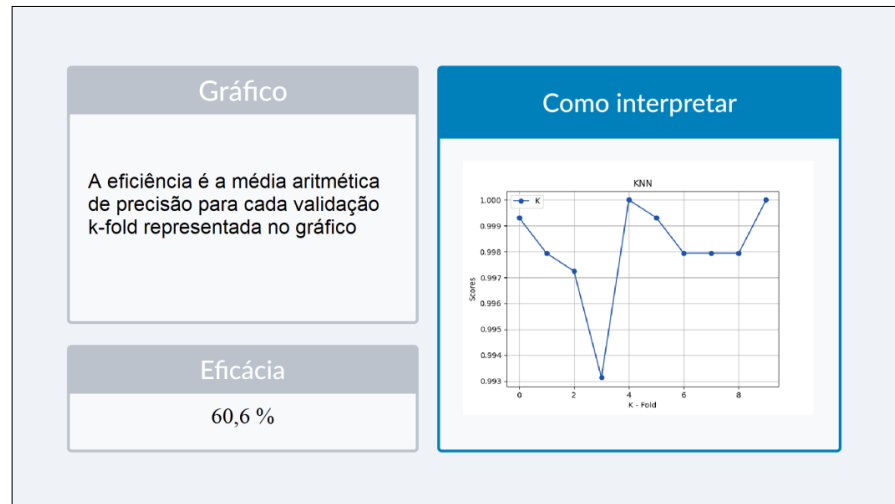


**Figura 2 ▼**  
Etapa de seleção do método de: (a) ML e (b) DL.  
Fonte: elaborado pelos autores



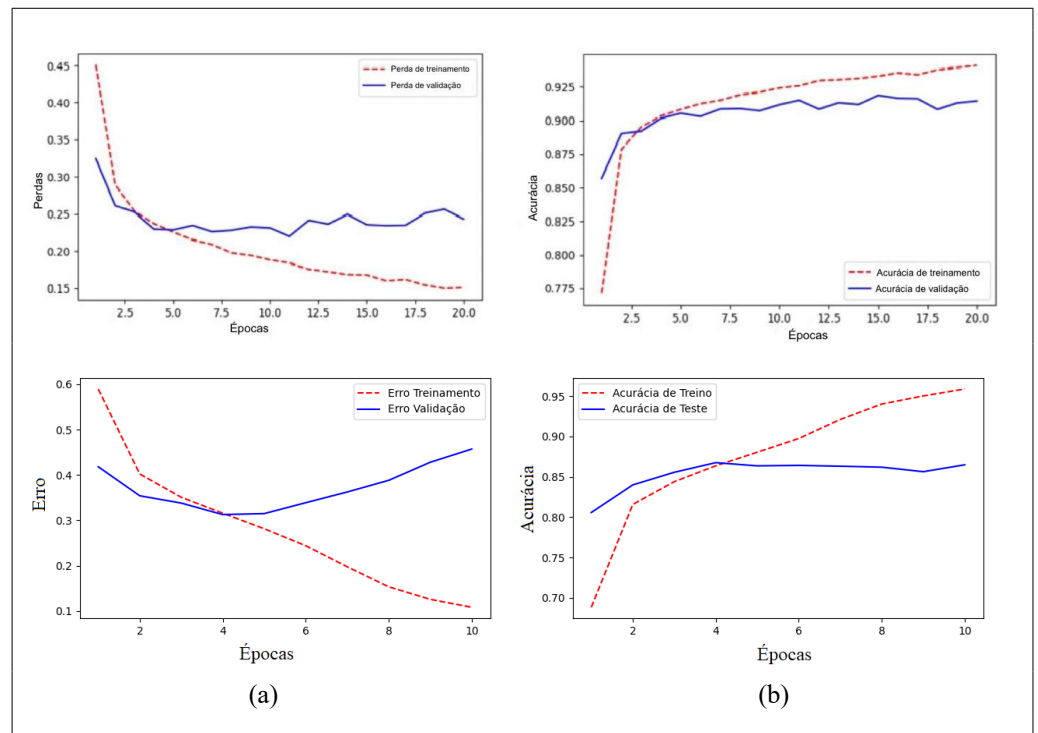
**Figura 3** ▶

Análise de resultados obtidos pelo Sim-ECG, utilizando a técnica de ML.  
 Fonte: elaborado pelos autores



**Figura 4** ▶

Análise de resultados obtidos pelo Sim-ECG, utilizando a técnica de DL: (a) erros do modelo, ou perdas de treinamento, e de validação em termos do número de épocas e (b) acurácias, ou assertividade, de treinamento e de validação em termos do número de épocas.  
 Fonte: elaborado pelos autores



### 4 Cenários de validação

Para validação do Sim-ECG, foram considerados quatro cenários. No primeiro cenário ( $C_1$ ), buscou-se implementar e validar os métodos propostos por Guvenir *et al.* (1997) –  $k$ NN, NB e VFI – para classificação de arritmias cardíacas. No segundo cenário ( $C_2$ ), foi aplicada uma CNN de uma dimensão (CNN-1D), proposta por Urtnasan *et al.* (2018), sendo esse método aplicado para detecção de anomalias presentes nos sinais de ECG e causadas pela apneia do sono. No terceiro cenário ( $C_3$ ), foi implementada a CNN proposta por Dey, Chaudhuri e Munshi (2018) que, de forma semelhante ao  $C_2$ , foi utilizada para classificação de anomalias presentes nos sinais de ECG e também causadas pela apneia do sono. No quarto cenário ( $C_4$ ), foram aplicados métodos que não possuíam bancos de dados pré-processados – a CNN proposta por Hsieh *et al.* (2020) e o SVM aplicado por Li *et al.* (2018) –, sendo esses utilizados apenas



como ferramentas para novas possibilidades de análise do Sim-ECG. No Quadro 2, descrevem-se os cenários utilizados para validação e testes do Sim-ECG.

**Quadro 2 ▶**

Cenários utilizados para validação e testes do Sim-ECG.  
Fonte: dados da pesquisa

Cenário	Método	Referência
C <sub>1</sub>	kNN, NB e VFI	Guvenir <i>et al.</i> (1997)
C <sub>2</sub>	CNN	Urtnasan <i>et al.</i> (2018)
C <sub>3</sub>	CNN	Dey, Chaudhuri e Munshi (2018)
C <sub>4</sub>	CNN e SVM	Hsieh <i>et al.</i> (2020) e Li <i>et al.</i> (2018)

Na Tabela 1, é possível observar os parâmetros de treinamento utilizados nas CNNs, considerando C<sub>2</sub> a C<sub>4</sub>. O primeiro cenário não consta na Tabela 1, pois o método de treinamento para os algoritmos de ML não é o mesmo que o utilizado para os algoritmos de DL. O método de treinamento adotado para os métodos de ML foi o de dobras, sendo escolhidas 10 dobras da respectiva base de dados.

**Tabela 1 ▶**

Parâmetros de treinamento utilizados nas CNNs, considerando C<sub>2</sub> a C<sub>4</sub>.  
Fonte: dados da pesquisa

Cenário	Amostras por iteração	Épocas
C <sub>2</sub>	128	20
C <sub>3</sub>	256	95
C <sub>4</sub>	256	10

Para obtenção dos resultados, foi utilizada a plataforma Google Colab, em que as especificações computacionais foram: placa de vídeo Tesla K80, com 2496 núcleos CUDA, 12 GB GDDR5 VRAM, processadores Xeon com um único núcleo e hiperencadeados a 2,3 GHz e memória RAM de 12 GB, 6 GB para cada simulação.

## 5 Resultados

Em C<sub>1</sub>, foi utilizada uma base de dados disponível na literatura (Guvenir *et al.*, 1997). Os dados foram separados em dez dobras, sendo que, em cada dobra, alterava-se a composição dos dados de treino e de teste. Para validação do Sim-ECG, os métodos analisados foram: kNN, NB e VFI.

Em comparação com os dados obtidos por Guvenir *et al.* (1997) – acurácia de 68,00% para o VFI, 50,00% para o kNN e 53,00% para o NB –, o Sim-ECG obteve 67,28%, 60,53% e 60,45%, resultando, aproximadamente, nos seguintes valores de erro relativo: 1,06%, 20,46% e 14,06%, respectivamente. Os valores para o kNN e o NB resultaram em maior erro relativo porque os algoritmos utilizados nos pacotes disponíveis na linguagem de programação Python são diferentes dos métodos utilizados por Guvenir *et al.* (1997). O menor erro relativo, referente ao método VFI, foi obtido após realização de ajustes em seus hiperparâmetros (estratégia e número de barras), visto que esses dados não foram explicitamente descritos por Guvenir *et al.* (1997). Em relação ao tempo de simulação, os algoritmos implementados (kNN e NB) não necessitaram de alto poder computacional, sendo necessários apenas alguns instantes para realização da simulação, enquanto o VFI, que teve seus hiperparâmetros ajustados, necessitou de aproximadamente 300 segundos

para realização da simulação. Após realizar dez simulações, considerando a mesma máquina utilizada, não houve variações de tempo nem de acurácia.

Em C<sub>2</sub>, separando-se a base de dados (Moody; Mark, 2000) em dados de treino e de teste – 80% e 20%, respectivamente –, utilizou-se o modelo de CNN proposto por Urtnasan *et al.* (2018). Esse modelo foi treinado por 20 épocas, considerando 128 amostras por iteração e utilizando o otimizador Adam (Bock; Weiß, 2019). A média da acurácia obtida pelo Sim-ECG foi de 91,43%, considerando dez simulações realizadas – enquanto Urtnasan *et al.* (2018) obtiveram o valor de 96,00% –, resultando em um erro relativo de aproximadamente 4,76%. Esse erro relativo foi devido ao fato de a disposição dos dados utilizada pelo Sim-ECG não ser a mesma utilizada no trabalho de Urtnasan *et al.* (2018), além de alguns parâmetros não terem sido explicitamente especificados. Em relação ao tempo de simulação, o algoritmo utilizado necessitou de um processamento mais elevado. Por esse motivo, o processamento paralelo, por meio da placa de vídeo, foi utilizado. Com isso, foram necessários, aproximadamente, 900 segundos por treinamento, e o tempo total de simulação foi de, aproximadamente, 2 horas e 30 minutos.

Em C<sub>3</sub>, foi considerada a mesma base de dados de C<sub>2</sub> e os mesmos valores dos dados de treino e de teste. No entanto, utilizou-se o modelo de CNN proposto por Dey, Chaudhuri e Munshi (2018), treinando-se o método por 95 épocas, com 256 amostras por iteração e utilizando o mesmo otimizador (Bock; Weiß, 2019). A média da acurácia obtida pelo Sim-ECG foi de 90,46%, considerando dez simulações realizadas – enquanto Dey, Chaudhuri e Munshi (2018) obtiveram o valor de 98,00% –, resultando em um erro relativo de aproximadamente 7,69%. Esse erro relativo se deve ao fato de diversos parâmetros não terem sido explicitamente descritos por Dey, Chaudhuri e Munshi (2018). O tempo de simulação para cada treinamento, nas mesmas condições de C<sub>2</sub>, foi de aproximadamente 1140 segundos, e o tempo total de simulação foi de, aproximadamente, 3 horas e 20 minutos. Na Tabela 2, apresentam-se os resultados obtidos pelo Sim-ECG, em termos de acurácia e tempo de simulação, em comparação com os dados disponíveis na literatura (Guvenir *et al.*, 1997; Dey; Chaudhuri; Munshi, 2018; Urtnasan *et al.*, 2018).

**Tabela 2 ▼**

Resultados obtidos, em termos de acurácia e tempo de simulação, em comparação com os dados disponíveis na literatura (Guvenir *et al.*, 1997; Dey; Chaudhuri; Munshi, 2018; Urtnasan *et al.*, 2018).  
Fonte: dados da pesquisa

Método	Acurácia (%) Referência	Acurácia (%) Sim-ECG	Erro relativo (%)	Tempo de simulação (s) Sim-ECG
VFI	68,00	67,28	1,06	300
NB	50,00	60,45	20,90	0,67
kNN	53,00	60,53	14,21	0,73
CNN	96,00	91,43	4,76	900
CNN	98,00	90,46	7,69	1140

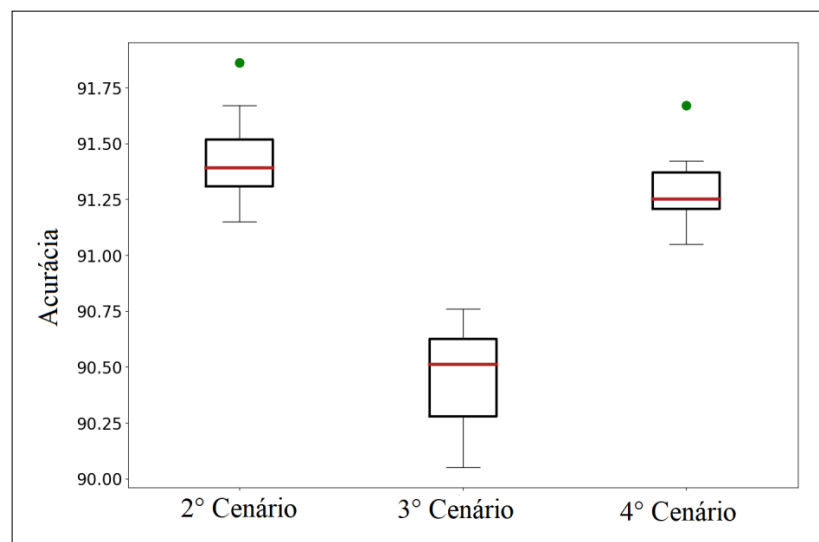
O cenário C<sub>4</sub> foi composto por métodos utilizados nos trabalhos de Hsieh *et al.* (2020) – CNN – e Li, Rajagopalan e Clifford (2013) – SVM –, de forma a utilizar o Sim-ECG com diferentes bases de dados. Para simulação da CNN proposta por Hsieh *et al.* (2020), foram consideradas as seguintes bases de dados: *apnea*-ECG (Moody; Mark, 2000) e *arritmia*-ECG (Moody; Mark, 2005). Não foi possível utilizar o banco de dados empregado por Guvenir *et al.* (1997), pois houve incompatibilidade de dimensões. Desse modo, foram utilizadas dez épocas e 256 amostras por iteração. O Sim-ECG obteve uma acurácia média de 91,29%, considerando dez simulações realizadas, com o tempo de simulação de, aproximadamente, 180 segundos para cada simulação e o tempo total de 30 minutos. Para o método SVM, utilizando a base de dados *arritmia*-ECG (Moody; Mark, 2005),



foram obtidos 100% de acurácia; já utilizando a base de dados de Guvenir *et al.* (1997), o resultado foi de 60,57%.

Na Figura 5, ilustra-se o diagrama de caixa de cada uma das simulações realizadas nos cenários C<sub>2</sub> a C<sub>4</sub>, envolvendo CNNs, em termos de acurácia. Os diagramas de caixa foram obtidos para um conjunto de dez simulações independentes. A caixa representa o intervalo entre o primeiro e o terceiro quartil, enquanto a linha horizontal dentro da caixa indica a mediana.

**Figura 5** ►  
Diagrama de caixa de cada uma das simulações realizadas nos cenários C<sub>2</sub> a C<sub>4</sub>, envolvendo CNNs, em termos de acurácia.  
Fonte: dados da pesquisa



## 6 Conclusões

Neste artigo, propôs-se o Sim-ECG, uma plataforma computacional de código aberto que analisa sinais obtidos por meio de exames de eletrocardiograma. Para isso, foram utilizadas bases de dados públicas que possibilitaram analisar e classificar as informações, considerando diversos métodos utilizados na área de inteligência artificial.

Ao realizar as etapas de extração de características e classificação dos batimentos cardíacos, o Sim-ECG obteve valor de acurácia muito próximo do resultado obtido por Guvenir *et al.* (1997) no primeiro cenário de simulação (aproximadamente 1,06% de erro relativo). Entretanto, no segundo e no terceiro cenários, foram obtidos maiores valores de erro relativo referente à acurácia (4,76% e 7,69%, respectivamente), devido ao fato de que alguns dados e parâmetros necessários para validação não foram explicitamente descritos nas referências analisadas.

Os resultados obtidos pelo Sim-ECG mostraram que a ferramenta proposta permite que diversas técnicas de inteligência artificial implementadas na literatura sejam replicadas, técnicas essas que possuem acurácia significativa para classificação de patologias cardíacas. Em especial, as CNNs, que, apesar de demandarem um poder de processamento computacional elevado, possuem uma assertividade considerável, além de apresentarem maior facilidade de processamento com maior volume de dados para treinamento, podendo assim gerar modelos de alta confiabilidade para aplicações reais.

Utilizando o Sim-ECG, é possível ainda construir protótipos de modelos, verificando a melhor performance para uma determinada base de dados, utilizando técnicas disponíveis de *machine learning* e *deep learning*. Uma vez que esses protótipos sejam testados, podem ser extraídos do código-fonte e aprimorados conforme a necessidade do usuário,

tendo em vista que o Sim-ECG possui código aberto e se encontra disponível para as comunidades acadêmica e científica.

## Agradecimentos

Os autores deste trabalho agradecem à UFCG o apoio institucional.

## Financiamento

Esta pesquisa não recebeu financiamento externo.

## Conflito de interesses

Os autores declaram não haver conflito de interesses.

## Contribuições ao artigo

**PEREIRA, H. A.; SILVA, J. P. S.:** concepção ou desenho do estudo/pesquisa; análise e/ou interpretação dos dados; revisão final com participação crítica e intelectual no manuscrito. **ALMEIDA, P. H. S.; CANANÉA, L. C. L.:** concepção ou desenho do estudo/pesquisa; revisão final com participação crítica e intelectual no manuscrito. Todos os autores participaram da escrita, discussão, leitura e aprovação da versão final do artigo.

## Referências

ASSEN, M. V.; MUSCOGIURI, G.; CARUSO, D.; LEE, S. J.; LAGHI, A.; DE CECCO, C. N. Artificial intelligence in cardiac radiology. **La Radiologia Medica**, v. 125, p. 1186-1199, 2020. DOI: <https://doi.org/10.1007/s11547-020-01277-w>.

BOCK, S.; WEIß, M. A proof of local convergence for the Adam optimizer. *In*: 2019 INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS (IJCNN), 2019, Budapest. **Proceedings** [...]. Budapest: IEEE, 2019, p. 1-8. DOI: <https://doi.org/10.1109/IJCNN.2019.8852239>.

BOUSSELJOT, R.-D. **Physionet PTB Diagnostic ECG Database**. 2004. Disponível em: <http://www.physionet.org/content/ptbdb/1.0.0>. Acesso em: 23 abr. 2023.

DEY, D.; CHAUDHURI, S.; MUNSHI, S. Obstructive sleep apnoea detection using convolutional neural network based deep learning framework. **Biomedical Engineering Letters**, v. 8, p. 95-100, 2018. DOI: <https://doi.org/10.1007/s13534-017-0055-y>.

FEENY, A. K.; CHUNG, M. K.; MADABHUSHI, A.; ATTIA, Z. I.; CIKES, M.; FIROUZNI, M.; FRIEDMAN, P. A.; KALSCHUR, M. M.; KAPA, S.; NARAYAN, S. M.; NOSEWORTHY, P. A.; PASSMAN, R. S.; PEREZ, M. V.; PETERS, N. S.; PICCINI, J. P.; TARAKJI, K. G.; THOMAS, S. A.; TRAYANOVA, N. A.; TURAKHIA, M. P.; WANG, P. J.

Artificial intelligence and machine learning in arrhythmias and cardiac electrophysiology. **Circulation: Arrhythmia and Electrophysiology**, v. 13, n. 8, e007952, 2020. DOI: <https://doi.org/10.1161/circep.119.007952>.

FERREIRA, J. L.; LOPES, F. G. **Eletrocardiograma para enfermeiros**. [S.l.]: BoD GmbH DE, 2013.

FLETCHER, G. S. **Epidemiologia clínica: elementos essenciais**. Porto Alegre: Artmed, 2021.

GUVENIR, H. A.; ACAR, B.; DEMIROZ, G.; CEKIN, A. A supervised machine learning algorithm for arrhythmia analysis. *In: COMPUTERS IN CARDIOLOGY, 1997, Lund. Proceedings [...]. Lund: IEEE, 1997. p. 433-436. DOI: <https://doi.org/10.1109/CIC.1997.647926>.*

HSIEH, C.-H.; LI, Y.-S.; HWANG, B.-J.; HSIAO, C.-H. Detection of atrial fibrillation using 1D convolutional neural network. **Sensors**, v. 20, n. 7, 2136, 2020. DOI: <https://doi.org/10.3390/s20072136>.

JIANG, B.; GUO, N.; GE, Y.; ZHANG, L.; OUDKERK, M.; XIE, X. Development and application of artificial intelligence in cardiac imaging. **British Journal of Radiology**, v. 93, n. 1113, 20190812, 2020. DOI: <https://doi.org/10.1259/bjr.20190812>.

KACHUEE, M.; FAZELI, S.; SARRAFZADEH, M. ECG heartbeat classification: a deep transferable representation. *In: IEEE INTERNATIONAL CONFERENCE ON HEALTHCARE INFORMATICS (ICHI), 2018, New York. Proceedings [...]. New York: IEEE, 2018. p. 443-444. DOI: <https://doi.org/10.1109/ICHI.2018.00092>.*

KILIC, A. Artificial intelligence and machine learning in cardiovascular health care. **The Annals of Thoracic Surgery**, v. 109, n. 5, p. 1323-1329, 2020. DOI: <https://doi.org/10.1016/j.athoracsur.2019.09.042>.

LI, Q.; RAJAGOPALAN, C.; CLIFFORD, G. D. Ventricular fibrillation and tachycardia classification using a machine learning approach. **IEEE Transactions on Biomedical Engineering**, v. 61, n. 6, p. 1607-1613, 2013. DOI: <https://doi.org/10.1109/TBME.2013.2275000>.

LI, Z.; DERKSEN, H.; GRYAK, J.; HOOSHMAND, M.; WOOD, A.; GHANBARI, H.; GUNARATNE, P.; NAJARIAN, K. Supraventricular tachycardia detection via machine learning algorithms. *In: IEEE INTERNATIONAL CONFERENCE ON BIOINFORMATICS AND BIOMEDICINE (BIBM), 2018, Madrid. Proceedings [...]. Madrid: IEEE, 2018. p. 2419-2422. DOI: <https://doi.org/10.1109/BIBM.2018.8621164>.*

MOODY, G.; MARK, R. **Physionet Apnea-ECG Database**. 2000. Disponível em: <http://physionet.org/content/apnea-ecg/1.0.0>. Acesso em: 23 abr. 2023.

MOODY, G.; MARK, R. **Physionet MIT-BIH Arrhythmia Database**. 2005. Disponível em: <http://www.physionet.org/content/mitdb/1.0.0>. Acesso em: 23 abr. 2023.

RAHMAN, Q. A.; TERESHCHENKO, L. G.; KONGKATONG, M.; ABRAHAM, T.; ABRAHAM, M. R.; SHATKAY, H. Utilizing ECG-based heartbeat classification for hypertrophic cardiomyopathy identification. **IEEE Transactions on Nanobioscience**, v. 14, n. 5, p. 505-512, 2015. DOI: <https://doi.org/10.1109/tnb.2015.2426213>.

SIONTIS, K. C.; NOSEWORTHY, P. A.; ATTIA, Z. I.; FRIEDMAN, P. A. Artificial intelligence-enhanced electrocardiography in cardiovascular disease management. **Nature Reviews Cardiology**, v. 18, n. 7, p. 465-478, 2021. DOI: <https://doi.org/10.1038/s41569-020-00503-2>.

SOBRAC – SOCIEDADE BRASILEIRA DE ARRITMIAS CARDÍACAS. **Perguntas e respostas sobre arritmias cardíacas**. 2022. Disponível em: <https://www.sobrac.org/campanha/perguntas-e-respostas-sobre-arritmias-cardiacas>. Acesso em: 23 abr. 2023.

URTNASAN, E.; PARK, J.-U.; JOO, E.-Y.; LEE, K.-J. Automated detection of obstructive sleep apnea events from a single-lead electrocardiogram using a convolutional neural network. **Journal of Medical Systems**, v. 42, 104, 2018. DOI: <https://doi.org/10.1007/s10916-018-0963-0>.