Journal of Health Informatics

J Health Inform. 2026;18:e1806
DOI: https://doi.org/10.59681/2175-4411.v18.iEspecial.2026.1806

ARTIGO COMPLETO
CBIS 2026

Cloud-based convolutional neural network for ECG arrhythmia classification

Rede neural convolucional baseada em nuvem para classificação de arritmias em ECG

Augusto Pretto Zatt[1], André Luís del Mestre Martins[2], Fábio Pires Itturriet[3], Juliano Costa Machado[2] ✉

ORCID: 0009-0006-0658-9321 · 0009-0008-2829-2494 · 0000-0003-3140-2259 · 0000-0002-9484-4145

[1] Universidade Federal do Rio Grande do Sul, Porto Alegre (RS), Brasil.
[2] Instituto Federal Sul-Riograndense, Charqueadas (RS), Brasil.
[3] Programa de Pós-Graduação em Engenharia Biomédica (PPGEB), Universidade Tecnológica Federal do Paraná (UTFPR), Curitiba (PR), Brasil.

ARTICLE INFORMATION

TYPE
Artigo completo · CBIS 2026

IDENTIFIER
e1806

DOI
10.59681/2175-4411.v18.iEspecial.2026.1806

RECEIVED
17 May 2026

ACCEPTED
20 July 2026

REVIEW PROCESS
Double Blind Review

HANDLING EDITOR
Juliano de Souza Gaspar

ACCESS
Open, CC BY 4.0

CORRESPONDING AUTHOR
Juliano Costa Machado
julianomachado@ifsul.edu.br

LICENSE

This is an open access article published under the Creative Commons CC BY 4.0 license, which permits use, distribution, and reproduction in any medium, provided the original work is properly cited.

CC BY 4.0

ABSTRACT

Objective: To deploy a pre-built deep learning model on a free cloud service, utilizing convolutional neural networks (CNNs) for the detection of cardiac rhythms in electrocardiogram (ECG) signals. **Method:** The MIT-BIH arrhythmia dataset was used, along with segmentation, normalization, and synthetic data generation techniques, to train a model created from convolutional networks and to test its operation in a cloud. The system tested the dataset with a digital filter and without a filter, and performance was evaluated using a confusion matrix and additional metrics. **Results:** The proposed model achieved an accuracy compatible with the current literature in classifying five different types of cardiac rhythms. An accuracy of 92% was obtained for the unfiltered dataset and 92.2% for the filtered dataset; for the F1-Score metric, 92.1% was achieved for the unfiltered dataset and 92.3% for the filtered dataset. **Conclusion:** The execution time and classification accuracy demonstrated satisfactory performance compared to current methods, making the investment in this type of technology worthwhile.

Keywords: Arrhythmias, Cardiac; Cloud Computing; Diagnosis, Computer-Assisted; Neural Networks, Computer.

RESUMO

Objetivo: Implantar um modelo de aprendizado profundo pré-construído em um serviço de nuvem gratuito, utilizando redes neurais convolucionais (CNNs) para a detecção de ritmos cardíacos em sinais de eletrocardiograma (ECG). **Método:** Foram usados o conjunto de dados MIT-BIH de arritmias e técnicas de segmentação, normalização e geração de dados sintéticos, para treinar um modelo baseado em redes convolucionais e testar sua operação na nuvem. O sistema testou o conjunto de dados com e sem filtro digital, e o desempenho foi avaliado por meio de uma matriz de confusão e métricas adicionais. **Resultados:** O modelo proposto alcançou uma acurácia compatível com a literatura atual na classificação de cinco diferentes tipos de ritmos cardíacos. Foi obtida uma acurácia de 92% para o conjunto de dados sem filtro e de 92,2% para o conjunto com filtro; para a métrica F1-Score, foram alcançados 92,1% para o conjunto sem filtro e 92,3% para o conjunto com filtro. **Conclusão:** O tempo de execução e a acurácia de classificação demonstraram desempenho satisfatório em comparação com os métodos atuais, o que justifica o investimento nesse tipo de tecnologia.

Descritores: Arritmias Cardíacas; Computação em Nuvem; Diagnóstico por Computador; Redes Neurais de Computação.

CBIS 2026 XXI CONGRESSO BRASILEIRO DE INFORMÁTICA EM SAÚDE

Paper presented at the 21st Brazilian Congress of Health Informatics (CBIS 2026), held from September 23 to 25, 2026, in Brasília (DF), Brazil, and published in the CBIS 2026 section of the Journal of Health Informatics.

HOW TO CITE THIS ARTICLE: Zatt AP, Martins ALM, Itturriet FP, Machado JC. Cloud-based convolutional neural network for ECG arrhythmia classification. J Health Inform 2026;18:e1806. doi: https://doi.org/10.59681/2175-4411.v18.iEspecial.2026.1806

1 INTRODUCTION

According to the World Health Organization (WHO), cardiovascular diseases (CVDs) remain the leading cause of global mortality, accounting for approximately 17.9 million deaths in 2019 [1]. Among the various manifestations of CVDs, this study focuses on electrical heart dysfunctions, particularly in the diagnosis of cardiac arrhythmias [2]. Cardiac arrhythmias can be classified according to their life-threatening risk and manifest as slow, rapid, or irregular heartbeats [3,4]. The electrocardiogram (ECG) is the most widely used diagnostic method for detecting these anomalies, generating signals that are subsequently analyzed morphologically by specialists. However, manual interpretation of the ECG can lead to variability in diagnoses, as different observers may reach different conclusions [5]. Arrhythmias that do not present an immediate life risk are classified into four main categories: non-ectopic, supraventricular ectopic, ventricular ectopic, and fusion [6]. Accuracy in classifying these arrhythmias is critical, as each requires distinct therapeutic approaches. However, manual evaluation of ECG signals is subject to human error and interpretation variability, highlighting the need for more objective and automated approaches.

With the growing evolution of digital technologies in healthcare, computer-aided diagnosis (CAD) systems have been widely explored to reduce subjectivity and the workload of professionals [5]. Traditionally, machine learning-based approaches for arrhythmia diagnosis involve steps such as preprocessing, feature extraction, dimensionality reduction, and classification. Although these methods can provide satisfactory results, the need to manually extract relevant features from ECG signals imposes a limitation, especially when applied to large volumes of data or new datasets [7]. The application of deep learning techniques has shown promise in addressing these limitations. Convolutional Neural Networks (CNNs) have proven effective in automatically extracting features directly from raw ECG signals [8]. Consequently, steps such as feature extraction and selection, traditionally required in machine learning, are implicitly performed by the model, increasing the efficiency of the classification process.

Comparative studies show that deep learning-based approaches outperform methods that use manual feature extraction combined with traditional classifiers, offering greater accuracy in arrhythmia classification [7]. Moreover, the advancement of cloud computing platforms such as Amazon Web Services (AWS) enables the efficient execution of these models on a large scale, allowing for massive ECG signal processing and reducing response time in diagnostics [9]. The integration of deep learning with cloud computing solutions is particularly relevant in the context of the Internet of Things (IoT). IoT devices, such as wearable ECG monitors, can capture and transmit data in real time to the cloud, where deep learning models can process these signals and generate automatic diagnoses [10]. This not only facilitates continuous patient monitoring but also extends the reach of healthcare, enabling accurate diagnoses to be made remotely, faster, and at a lower cost. The scalability and flexibility provided by cloud computing make this approach essential for the development of large-scale CAD systems [11].

In the work [12], a hybrid approach is implemented, combining CNNs and long short-term memory (LSTM), using features extracted from the Continuous Wavelet Transform (CWT) with one-dimensional CNNs. In this case, cloud processing occurs at the edge, encompassing steps from feature extraction to data classification. Although the work describes the use of a quantization algorithm to reduce transmission volume and optimize latency without significantly compromising performance, it does not provide specific analyses regarding latency or resource consumption, focusing instead on reporting the system's implementation on a Raspberry Pi device.

Article [13] discusses the development and implementation of a cloud-based artificial intelligence system for the diagnosis of cardiac arrhythmias, utilizing a recurrent neural network with LSTM units. The solution aims to enable continuous monitoring of ECG signals by remotely processing the stages of filtering, segmentation, feature extraction, and heartbeat classification. Cloud computing offers advantages such as increased processing power and reduced energy consumption for wearable devices. The system executed on an AWS virtual machine highlights technical limitations such as latency and a processing capacity of up to 96 beats per minute, while proposing future optimizations to meet the demands of continuous monitoring.

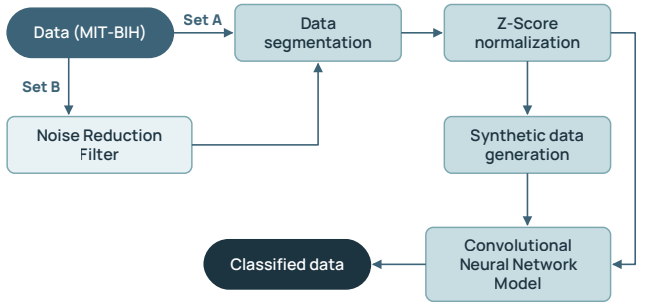
Thus, this work proposes the development of a deep learning model based on CNNs for the automatic classification of cardiac arrhythmias from ECG signals. The system is designed to be scalable and efficient, leveraging AWS cloud infrastructure to process large volumes of ECG data. The combination of these technologies, along with IoT integration, has the potential to make arrhythmia diagnosis more accessible, accurate, and less dependent on human interpretation.

2 METHODS

This work utilizes a previously constructed CNN based on the study by Acharya U. R. *et al.* [7]. With the network in place, the model was trained, according to Figure 1, and deployed on a free cloud service. For testing purposes, two networks were trained: one using a high-pass filter in cascade with a notch filter, and the

other without any filtering applied, using the raw ECG signal. In this section, we present the methodology step by step.

Figure 1. Model training flow



Source: Prepared by the authors (2026).

2.1 Database and Preprocessing

The dataset used was the MIT-BIH Arrhythmia Database [14], comprising annotated ECG signals from 47 patient records, where 25 presented cardiac arrhythmias and 22 were normal. This study focused on the MLII (Modified Lead II) lead, which offers high-quality signals and is widely used in arrhythmia classification. To ensure model robustness in scenarios with and without preprocessing, the study used two datasets:

- Set A: composed of raw ECG signals.
- Set B: processed through a digital filter cascade, consisting of a third-order high-pass filter to attenuate frequencies below 0.5 Hz, followed by a notch filter at 60 Hz to eliminate power line interference [14]. The transfer function of these filters was defined to retain frequencies relevant to ECG signals and minimize external noise, with a Bode plot illustrating the effectiveness of these cutoffs. The transfer functions are given by equations (a) and (b), and their Bode plots are shown in Figure 2 and Figure 3, respectively.

$$H(z) = \frac{0.99 - 2.97z^{-1} + 2.97z^{-2} - 0.99z^{-3}}{1 - 2.98z^{-1} + 2.961z^{-2} - 0.98z^{-3}} \quad \text{(a)}$$

$$H(z) = \frac{1 - z^{-1} + z^{-2}}{1 - 0.995z^{-1} + 0.99z^{-2}} \quad \text{(b)}$$

2.2 Segmentation and Normalization

The segmentation phase organized the signals into windows of 260 samples, centered on the R-peak of each heartbeat. This process standardizes input data, allowing the model to learn features from signals with uniform shapes. To minimize amplitude and offset variations, Z-score normalization was applied, adjusting each segment's values to zero mean and unit standard deviation, according to equation (c). This technique improves data consistency, as suggested by Acharya *et al.* [7]. Figure 4 demonstrates how the signal is segmented.

$$Z = \frac{X - \mu}{\sigma} \quad \text{(c)}$$

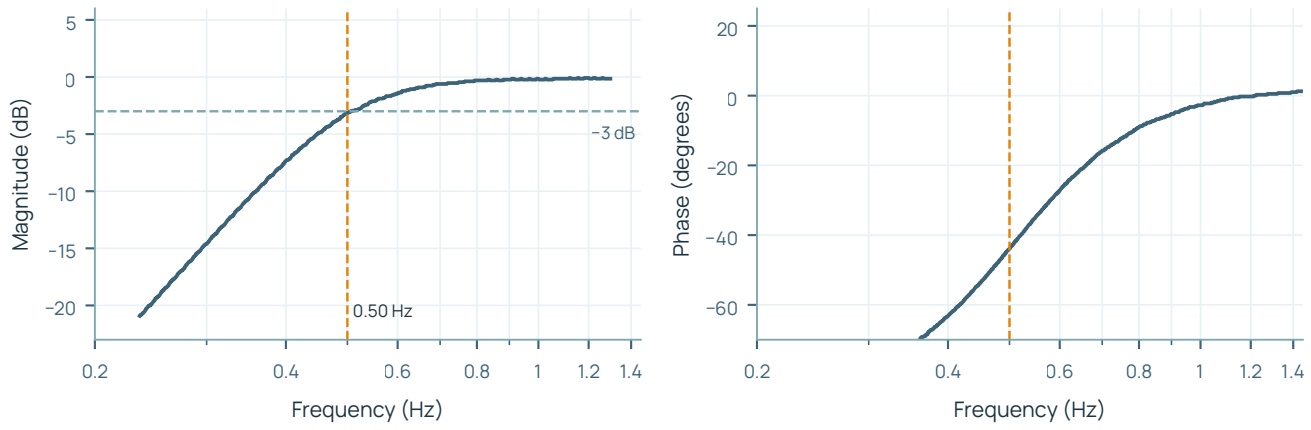
Where:

- Z is the normalized value;
- X is the original value of the signal;
- μ is the mean of the signal values;
- σ is the standard deviation of the signal values.

2.3 Synthetic Data Generation

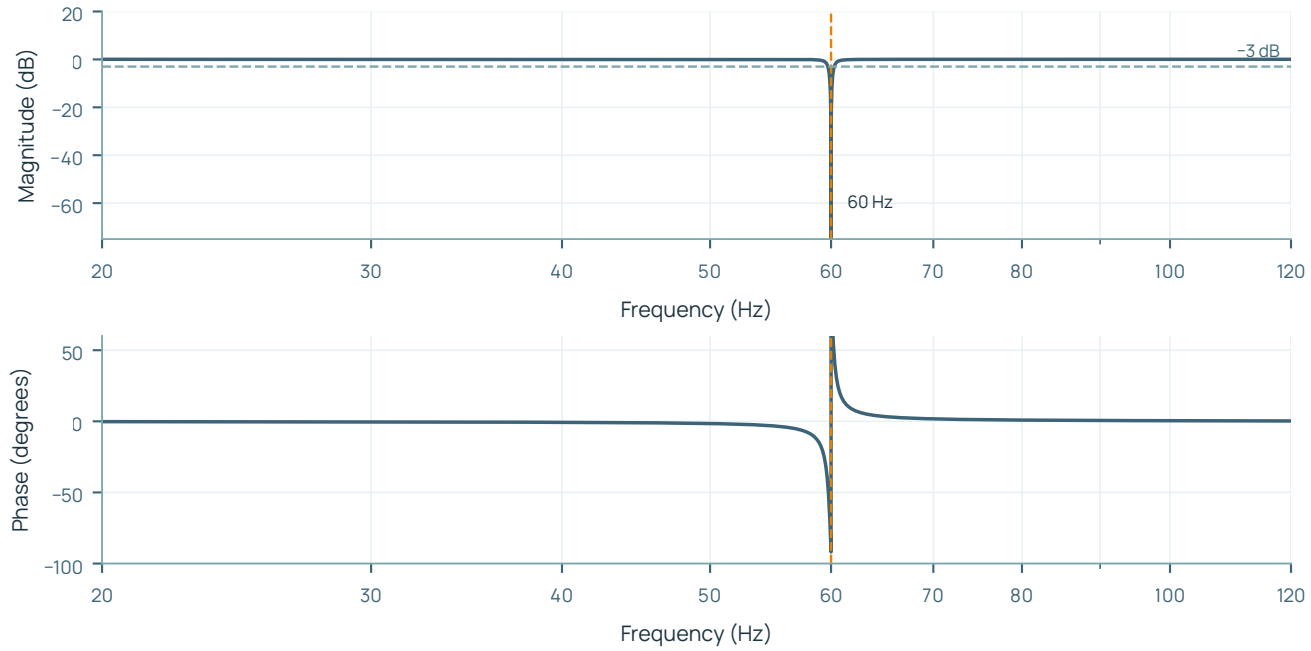
To balance the number of examples among different arrhythmia classes, a synthetic data generation technique was implemented. Initially, the original segments were “denormalized” by multiplying by the standard deviation and adding the mean. Then, slight variations of up to 10% in mean and standard deviation were applied, creating new examples without altering the essential signal morphology [7]. Figure 5 shows the comparison between the original signal and the synthetic signal generated from it.

Figure 2. Bode plot: high-pass filter



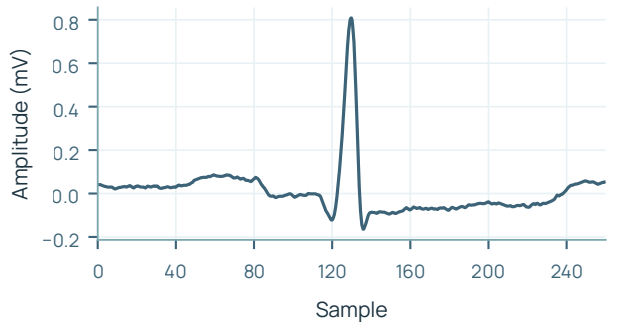
Source: Prepared by the authors (2026).

Figure 3. Bode plot: 60 Hz notch filter



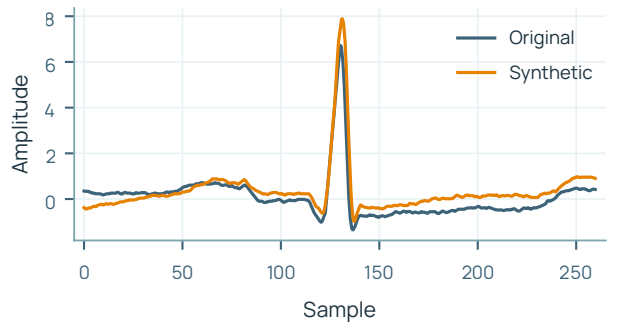
Source: Prepared by the authors (2026).

Figure 4. Signal segmentation



Source: Prepared by the authors (2026).

Figure 5. Comparison between the original signal and the generated synthetic signal



Source: Prepared by the authors (2026).

2.4 Convolutional Neural Network Architecture

The developed CNN follows a structure inspired by Acharya *et al.* [7], with three convolutional layers alternated by pooling layers and concluded with two dense layers. The first convolutional layer applies five filters with a kernel size of 3, while the subsequent layers apply 10 and 20 filters with kernel sizes of 4. Each layer uses the ReLU activation function and L2 regularization with a value of 0.02 to reduce overfitting. The model was compiled with the categorical_crossentropy loss function and optimized with the Adam algorithm, as validated by Bouvrie [15]. The final output layer uses softmax, producing probabilities for the five arrhythmia classes.

2.5 Training and Validation

Model training was conducted over 20 epochs with tenfold cross-validation. The entire ECG dataset was first separated into class-specific vectors, and ten data folds were subsequently created by randomly selecting signals from each class vector, ensuring that all heartbeat classes were represented in every fold. These ten folds were then used to split the dataset into ten equal parts, as recommended by Duda and Hart [16]. In each iteration, nine data parts were used for training, and one part was reserved for validation. This method ensured stable model performance with results that do not depend on a single data split. Metrics used included recall, precision, and F1-score, calculated for each class in a one-vs-all approach.

2.6 Implementation on AWS Cloud

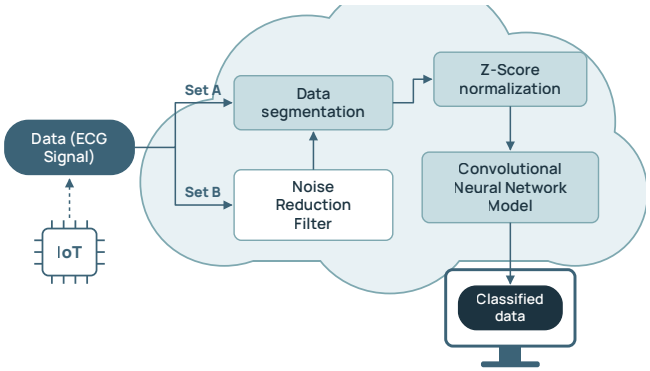
To explore the feasibility of a cloud-based diagnostic system, the model was implemented on an AWS environment, utilizing the platform's free-tier GPU support. This setup allowed testing the model's

scalability in a cloud service and evaluating real-time response for cardiac arrhythmia detection, essential for clinical diagnosis. Adopting AWS also enabled the simulation of a practical application where the neural network processes ECG signals in a production environment.

For this application, the parameters of the trained models were allocated in the cloud, and then Python code was implemented in the cloud to enable the classifier to be used with a test set. For this test, a signal with approximately 5 minutes of electrocardiogram recordings was used. For the test with filtering, the algorithm receives the signal, applies the filter, normalizes, segments, and then sends this signal to the model, which uses the previously trained parameters to classify the signal. For the test without filtering, the filtering step is omitted, but all other steps are maintained.

The algorithm's input consists solely of the electrocardiogram signal, and when combined with an R-peak classifier, it becomes a simple alternative for beat classification, requiring only an electrocardiogram signal with a minimum of 130 samples on each side of the peak as input, which can be extracted from an IoT device. The functioning of the algorithm is demonstrated in Figure 6.

Figure 6. Functioning of the algorithm



Source: Prepared by the authors (2026).

3 RESULTS

For both datasets, training was conducted in the Google Colaboratory (Google Colab) environment and lasted over 9,500 seconds. The developed algorithm required a high-performance GPU and fast RAM, given the model's complexity and data volume.

Table 1 shows the comparison between the metrics obtained with the two datasets. Precision, recall, and F1-score are computed using a one-vs-rest approach and averaged across classes:

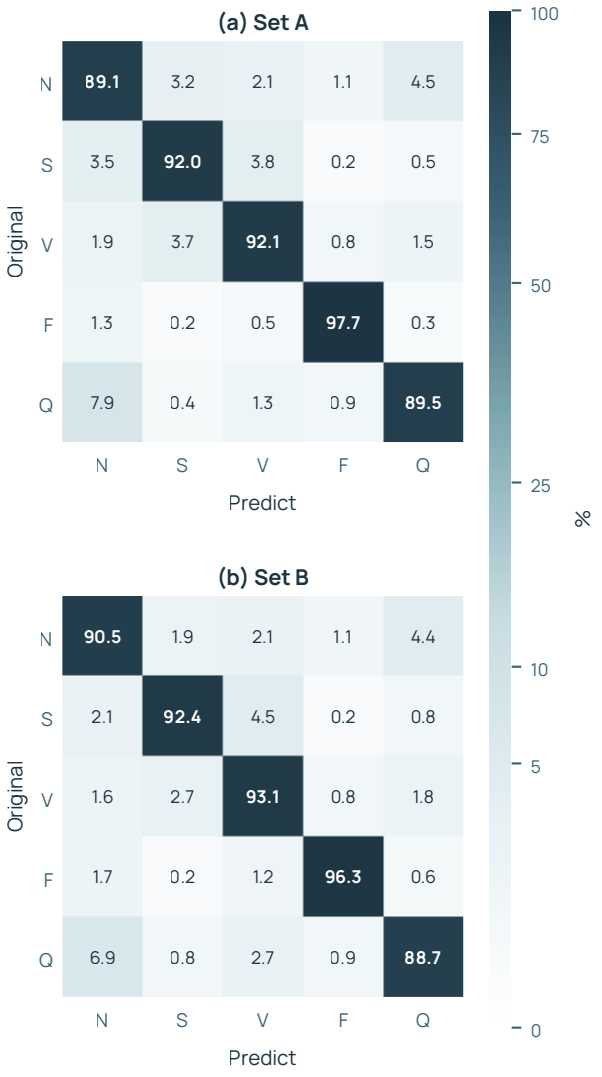
Table 1. Comparison of metrics

Set	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
A	92.0	92.1	92.1	92.1
B	92.2	92.4	92.2	92.3

Source: Prepared by the authors (2026).

In Figure 7, the confusion matrices for each dataset are presented in percentage.

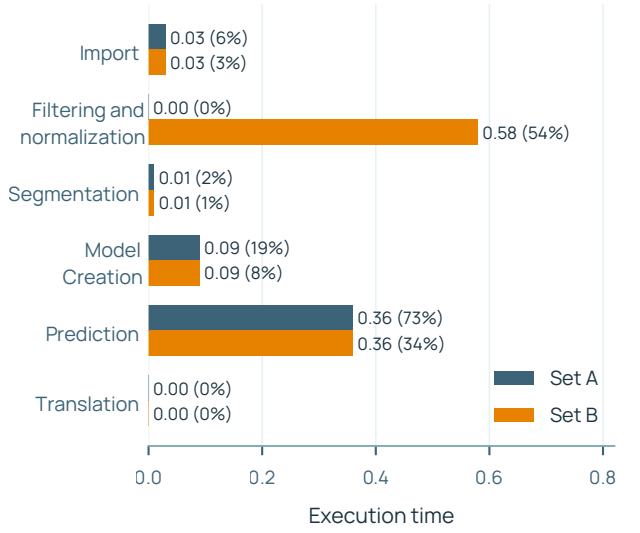
Figure 7. Confusion Matrix for set A (a) and set B (b) (%)



Source: Prepared by the authors (2026).

Figure 8 summarizes the execution time of each stage of the algorithm, including data import, filtering and normalization, segmentation, model inference, classification, and mapping of outputs to one of the five classes, for each dataset. In Table 2, different model configurations were used to evaluate the average execution time for each.

Figure 8. Execution times



Source: Prepared by the authors (2026).

Table 2. Average execution time (seconds) for different configurations.

Model	5 minutes of ECG signal	1 minute of ECG signal
Filter and the R-peaks provided	15.14	13.20
Filter and the R-peaks calculated	16.01	14.40
Without filter and the R-peaks provided	14.71	14.34
Without filter and the R-peaks calculated	15.64	14.51

Source: Prepared by the authors (2026).

4 DISCUSSION

Filtering Set B provided slight improvements in accuracy and precision. Filtering contributed to reducing false negatives, which is critical in arrhythmia detection.

The confusion matrices show that approximately 8% of samples were misclassified, indicating strong classification performance. The similarity between filtered and unfiltered results suggests that a CNN can learn to distinguish relevant signal patterns even in the presence of noise.

The training time exceeding 9,500 seconds reflects the computational cost of deep learning models, reinforcing the need for high-performance hardware. However, once trained, the model demonstrated fast and efficient prediction, supporting real-time applications.

The use of synthetic data was essential to address class imbalance and improve model robustness. However, because the synthetic samples were generated from the original dataset, they may have limited the generalizability of the reported performance, and the results should therefore be interpreted within the context of the augmented dataset. Additionally, the use of 10-fold cross-validation ensured stable and reliable performance.

Compared to the study by Acharya *et al.* [7], which reported higher accuracies (93.47% unfiltered and 94.03% filtered), the results obtained here are slightly lower. However, the consistency between filtered and unfiltered performance in both studies reinforces the robustness of CNN-based approaches.

A key advantage of this model is the elimination of manual feature extraction, simplifying the pipeline and increasing applicability in real-world scenarios. The ability to operate effectively even without filtering is particularly valuable for practical deployments.

Regarding execution time, filtering substantially increases processing time in the virtual machine scenario, accounting for approximately 54% of total execution time. However, in cloud-based scenarios, the difference between filtered and unfiltered execution times becomes negligible, making filtering a viable option when higher accuracy is desired.

The results demonstrate that CNNs are highly effective for ECG signal classification and can be integrated into cloud-based healthcare systems. The model's ability to operate using only raw ECG signals further enhances its applicability in IoT and wearable device contexts.

A limitation of the model is the requirement of at least 130 samples before and after an R-peak for segmentation. However, this is generally not restrictive in practical ECG acquisition scenarios.

Overall, the proposed system demonstrates strong potential for real-world application, combining accuracy, robustness, and computational feasibility.

5 CONCLUSION

This work implemented a heartbeat classifier based on CNN in the free AWS cloud and presented classification accuracy rates, in addition to exploring the system's execution data. It was also possible to compare the execution speed of a dedicated virtual machine with that of a healthcare ecosystem, bringing the tests closer to a potential application in medical routines. The execution time and classification accuracy demonstrated satisfactory performance compared to current methods, making the investment in this type of technology worthwhile. An investment in a more robust cloud environment could further improve performance, and with the promising market prospects for this technology, future work may focus on optimizing model parameters and investigating alternative data balancing techniques to further enhance generalization and classification performance.

Author contributions. Augusto Pretto Zatt implemented all algorithms and conducted the experiments. André Luís del Mestre Martins supervised the implementation of the algorithms in the cloud. Fabio Pires Itturriet contributed to the experimental design and the analysis of the results. Juliano Costa Machado coordinated the study and defined its methodology.

Funding. The authors did not declare any source of funding.

Data and materials availability. The source code for the algorithms used in this study is available from the corresponding author upon request.

Conflicts of interest. The authors did not declare any conflicts of interest.

Use of artificial intelligence. Artificial Intelligence tools were used to support text writing and revision (clarity, grammar, and style), as well as the organization and structuring of the manuscript.

REFERENCES

1. World Health Organization. Cardiovascular diseases (CVDs) [Internet]. Geneva: World Health Organization; 2021 [cited 2026 Mar 26]. Available from: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)
2. Al-Khatib SM, Stevenson WG, Ackerman MJ, Bryant WJ, Callans DJ, Curtis AB, et al. 2017 AHA/ACC/HRS Guideline for Management of Patients With Ventricular Arrhythmias and the Prevention of Sudden Cardiac Death: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Rhythm Society. Circulation. 2018;138(13):e272-e391. doi: https://doi.org/10.1161/CIR.0000000000000549
3. Acharya UR, Krishnan SM, Spaan JA, Suri JS, editors. Advances in Cardiac Signal Processing. Berlin, Heidelberg: Springer Berlin Heidelberg; 2007. doi: https://doi.org/10.1007/978-3-540-36675-1
4. National Heart, Lung, and Blood Institute. Types of arrhythmias [Internet]. 2022 [cited 2026 Oct 18]. Available from: https://www.nhlbi.nih.gov/health/arrhythmias/types
5. Martis RJ, Acharya UR, Adeli H. Current methods in electrocardiogram characterization. Comput Biol Med. 2014;48:133-49. doi: https://doi.org/10.1016/j.compbiomed.2014.02.012
6. Association for the Advancement of Medical Instrumentation. ANSI/AAMI EC57:2012/(R)2020. Testing and reporting performance results of cardiac rhythm and ST segment measurement algorithms. Arlington (VA): AAMI; 2020.
7. Acharya UR, Oh SL, Hagiwara Y, Tan JH, Adam M, Gertych A, et al. A deep convolutional neural network model to classify heartbeats. Comput Biol Med. 2017;89:389-96. doi: https://doi.org/10.1016/j.compbiomed.2017.08.022
8. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436-44. doi: https://doi.org/10.1038/nature14539
9. Goodfellow I, Bengio Y, Courville A. Deep learning. Cambridge (MA): MIT Press; 2016.
10. Khelifi H, Luo S, Nour B, Sellami A, Moungla H, Ahmed SH, et al. Bringing Deep Learning at the Edge of Information-Centric Internet of Things. IEEE Commun. Lett. 2019;23(1):52-5. doi: https://doi.org/10.1109/lcomm.2018.2875978
11. Mansour RF, Amraoui AE, Nouaouri I, Diaz VG, Gupta D, Kumar S. Artificial Intelligence and Internet of Things Enabled Disease Diagnosis Model for Smart Healthcare Systems. IEEE Access. 2021;9:45137-46. doi: https://doi.org/10.1109/access.2021.3066365
12. Han Y, Han P, Yuan B, Zhang Z, Liu L, Panneerselvam J. Novel Transformation Deep Learning Model for Electrocardiogram Classification and Arrhythmia Detection using Edge Computing. J Grid Computing. 2024;22(1):7. doi: https://doi.org/10.1007/s10723-023-09717-3
13. Freire MA, Itturriet FP, Del Mestre Martins AL, Abreu JP, Machado JC. Aprendizagem de máquinas com computação na nuvem para detecção de arritmias cardíacas. In: Anais do XXIX Congresso Brasileiro de Engenharia Biomédica (CBEB) [Internet]. Ribeirão Preto: Sociedade Brasileira de Engenharia Biomédica; 2024. p. 164-9. Disponível em: https://sbeb.org.br/wp-content/uploads/2024/10/SBEB-CBEB2024-ProceedingsFinal.pdf
14. Aboutabikh K, Aboukerdah N. Design and implementation of a multiband digital filter using FPGA to extract the ECG signal in the presence of different interference signals. Comput Biol Med. 2015;62:1-13. doi: https://doi.org/10.1016/j.compbiomed.2015.03.034
15. Bouvrie J. Notes on convolutional neural networks. Cambridge (MA): Massachusetts Institute of Technology; 2006.
16. Duda RO, Hart PE, Stork DG. Pattern classification. 2nd ed. New York: John Wiley & Sons; 2001.

EDITORIAL NOTES

The Journal of Health Informatics is a scientific journal maintained by the Brazilian Society of Health Informatics (Sociedade Brasileira de Informática em Saúde), a non-profit scientific association dedicated to the development of Health Informatics and Digital Health in Brazil. The editorial process is conducted with scientific independence, free from institutional, commercial, or sponsor interference, and follows international best practices in publication ethics.

This article was double-blind peer reviewed by specialists in the field. The editorial team used artificial intelligence tools, under human supervision, in the initial screening, in checking the text, references and numbers and, when needed, in redrawing figures, charts and tables, which the authors approved in the proof. The tools and safeguards are described in the journal's Policy on the use of artificial intelligence. Editorial decisions were made by people.

SBIS Sociedade Brasileira de Informática em Saúde