# Extraction of Signal Features for Diagnosing Heart Disease Using a Single Derivation of an Electrocardiogram

# Extração de Características de Sinais para Diagnóstico de Doença Cardíaca Utilizando Uma Única Derivação de um Eletrocardiograma

Evaldo Renó Faria Cintra[1], Gabriel Antonio Fanelli de Souza[2]

1 - Pós Doutorando na Universidade Federal de Itajubá - UNIFEI, Professor no Centro Universitário de Itajubá – FEPI. E-mail.: evaldocintraposdoc@unifei.edu.br
2 - Doutor em Engenharia Eletrônica e Computação no Instituto Tecnológico de Aeronáutica – ITA. Professor Adjunto da Universidade Federal de Itajubá – UNIFEI.E-mail.: fanelli@unifei.edu.br

Recebido em: 08/11/2024 Revisado em: 30/06/2025 Aprovado em: 07/07/2025

**Abstract:** In this study, we present a signal processing procedure that identifies characteristics of electrocardiogram signals and feeds this information into a neural network, enabling the separation of normal signals from those indicative of infarction. The method is tailored for integrating with neural network systems designed to diagnose based on signal characteristics. Using the proposed features, it was possible to differentiate normal signals from infarction signals, using only one derivation of the electrocardiogram. The proposed method aims to simplify the process of acquiring and processing electrocardiogram (ECG) signals. It involves extracting specific characteristics from an ECG signal, which are then provided to an intelligent system for potential diagnosis. Specifically, the method focuses on utilizing the V6 derivation of ECG signal. During processing, two key features are extracted: the average energy of the signal and the dominant frequency in the signal's frequency spectrum. To validate the effectiveness of this approach, a dataset consisting of 62 normal ECG signals and 58 signals diagnosed with infarction signals were employed to train the network and a further 480 signals were used for testing. Both the training and validation datasets were obtained from the PhysioNet signal bank. Each signal's extracted features were utilized as inputs for the neural network. The results showed that the network achieved a correct diagnosis rate of 99,79% using these signal characteristics.
**Key-words:** Heart. Neural Network. Derivation. Diagnoses. Electrocardiogram.

**Resumo:** Neste estudo, apresentamos um procedimento de processamento de sinais que identifica características dos sinais do eletrocardiograma e alimenta essa informação em uma rede neural, possibilitando a separação dos sinais normais daqueles indicativos de infarto. O método é adaptado para integração com sistemas de redes neurais projetados para diagnóstico com base nas características do sinal. Utilizando os recursos propostos, foi possível diferenciar sinais normais de sinais de infarto, utilizando apenas uma derivação do eletrocardiograma. O método proposto visa simplificar o processo de aquisição e processamento de sinais de eletrocardiograma (ECG). Envolve a extração de características específicas de um sinal de ECG, que são então fornecidas a um sistema inteligente para diagnóstico potencial. Especificamente, o método concentra-se na utilização da derivação V6 do sinal de ECG. Durante o processamento, duas características principais são extraídas: a energia média do sinal e a frequência dominante no espectro de frequência do sinal. Para validar a eficácia desta abordagem, um conjunto de dados, composto por 62 sinais normais de ECG e 58 sinais diagnosticados como sinais de infarto foram utilizados para treinar a rede e para o teste foram empregados mais 480 sinais. Os conjuntos de dados para treinamento e validação foram obtidos do banco de sinais PhysioNet. As características extraídas de cada sinal foram utilizadas como entradas para a rede neural. Os resultados mostraram que a rede alcançou uma taxa de diagnóstico correto de 99,79% utilizando essas características de sinal.
**Palavras-chave:** Coração. Rede Neural. Derivação. Diagnóstico. Eletrocardiograma.

## Introduction

The electrocardiogram signal is one of the most significant tools for identifying characteristics that make it possible to detect heart problems. According to the World Heart Federation, cardiovascular diseases were responsible for 20.5 million deaths in 2021, almost one-third of all deaths worldwide (Di Cesare et al., 2023).

Early diagnosis can often prevent these deaths. Many tools allow an early diagnosis of heart disease, however, those that can help diagnose and treat diseases are not accessible in the communities that need them most. Much research has been carried out looking for tools that allow the detection of cardiac anomalies (Gupta et al, 2022, Patel et al, 2021, Jager; Moody; Mark, 1998).

Currently, extensive research is focused on applying artificial intelligence to assist with decision-making applications. Artificial Intelligence (AI) is used in a wide variety of applications that involve learning, perception, and specific tasks such as chess games, demonstrating mathematical theorems, creating poetry, driving a car on the road, and disease diagnosis (Rahman et al., 2024; Alkhodari; Fraiwan, 2021). Neural networks represent a powerful subset of AI models adept at addressing diverse challenges across fields such as control, pattern recognition, and decision-making. Typically, these networks find their implementation primarily in software, using microprocessors for both training and execution phases (Shen et al., 2020; Dally et al., 2018). In the domain of pattern recognition, neural networks play a crucial role in discerning typical patterns, such as those found in electrocardiograms (ECGs).

Numerous studies have explored various approaches to detecting cardiac irregularities. Gupta proposed Chaos theory together with short-time Fourier transform (STFT) that reduces the occurrence of spurious outcomes and has demonstrated the properties that are typical outlook of deterministic chaotic systems. (Gupta; Mittal; Mittal, 2020). Maglaveras present a supervised neural network-based algorithm used for the automatic detection of ischemic episodes resulting from ST segment elevation or depression. For this leads V3, V4, V5, and MLIII were used (Maglaveras N. et al., 1998). Jager propose an evaluation and performance measurement protocol for database application. A method was proposed to evaluate the accuracy in detecting ST segment deviations, using two leads (Jager et al., 1991). Anuradha presented a technique that applies the extraction of 4 parameters from an ECG signal in lead II and, through a fuzzy classifier, identifies possible cardiac anomalies (Anuradha & Reddy, 2008). Zheng proposed a new improved QRS detector that uses Discrete Wavelet Transform and Cubic Spline Interpolation for preprocessing, together with a dynamic weight adjustment strategy to enhance the detection robustness in noisy conditions. In this application, derivation II was used. (Zheng & Wu, 2008). Sharma presented a novel technique for QRS complex detection with minimum pre-processing required.

The technique uses two-stage median filter and Savitzky-Golay smoothing filter for pre-processing. The decision rule is based on root-mean-square of the signal. Using one ECG channels (Sharma & Sunkaria, 2016). Chen presented a classification system based on wavelet transformation and probabilistic neural network to identify six ECG beat types (Chen & Yu, 2006). In this application, one derivation was considered for analysis.

Afsar proposed an efficient data pruning algorithm incorporated into a fuzzy k-nearest neighbor classifier to minimize space and time complexity during classification (Afsar et al., 2008). This work seeks to offer a novel solution: the early detection of cardiac issues using only a single derivation of the ECG signal.

With the use of only one derivation to determine a possible diagnosis of a patient, the equipment used would be simpler as it would require fewer points of contact with the patient's skin to acquire the signal. Furthermore, this proposal provides simpler processing for extracting characteristics from the ECG signal as it works with two input values from the network. This system enables the development of streamlined portable equipment capable of delivering diagnoses without requiring numerous connections between the patient and the apparatus. Through the adoption of such techniques, automated patient diagnosis becomes feasible, facilitating remote monitoring. Consequently real-time triggering of alarms can alert either the patient or medical personnel to take preemptive measures, mitigating the risk of addressing issues only at an advanced stage.

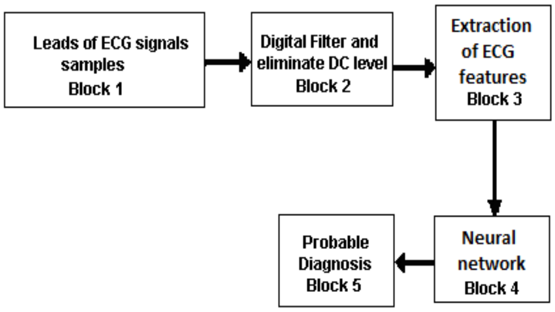
In this study, a solution is proposed using a neural network consisting of 50 neurons with two inputs and one output. Each input receives a characteristic, the network is able to generate a probable diagnosis of the patient under analysis. Our objective was to present a signal processing solution that uses two characteristics of a single derivation of the electrocardiogram signal to diagnose the presence of a heart attack. In this solution, the electrocardiogram signals from a PhysioNet database were used.

With the information in its inputs, the neural network performs the processing and informs in its output the possible diagnosis of the analyzed signal, whether it is a characteristic sign of a heart attack.

## Material and Methods

The proposed system was implemented in MATLAB®, and the electrocardiogram signals used to make the database and perform the test were obtained from the PhysioNet database (PhysioNet PhysioBank, 2023). After the simulation, to validate the result obtained with other works in the literature, figures of merit were calculated. Figure 1 shows a summary of the validation system used.

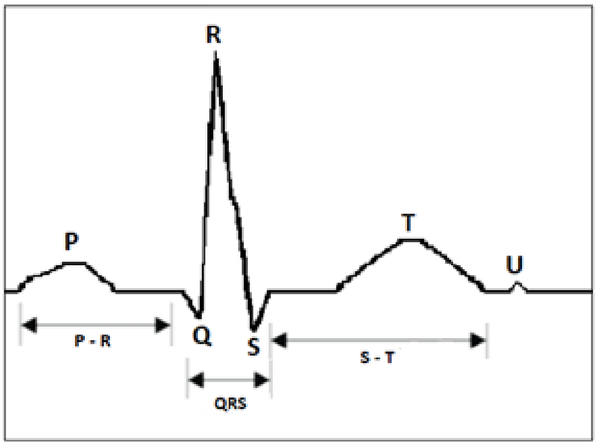
Figure 1. Describing the validation system




Initially, as shown in the first block, we obtained the leads of the ECG signals. The electrocardiogram is a test usually performed to assess cardiac rhythm disturbances. Using the electrocardiogram, one can obtain information on structural cardiac problems, e.g. myocardial ischemia, myocardial electrophysiological disturbances, pericardial diseases, heart position, cardiac pacing, and electrolytic and metabolic changes in the system. Additionally, it serves as a valuable tool for documenting both intrinsic and pharmacologically-induced alterations in cardiac function (Hampton & Hampton, 2019). The signal obtained in this test describes the entire cycle of the heart conduction system. Figure 2 shows a typical, normal ECG wave.

Observing Figure 2, it is possible to notice the main gaps highlighted. Dignosis of a cardiopathy can be made by assessing and detecting any amplitude variation or time variation during each interval.

Figure 2. Typical, normal ECG wave with its segments indicated.




In this system, each signal is represented by 10,000 samples at a sampling frequency of 1KHz.

In the second block the sampled signal is digitally filtered using a third-order Butterworth low-pass filter. The key characteristic of this filter is a ripple-free frequency response within its bandwidth and zero response outside its bandwidth. The magnitude response of an Nth-order Butterworth filter with a cutoff frequency wp is given by Equation. (1) (SEDRA, 2000).

$$|T(jw)| = 1/\sqrt{1 + exp^2(w/w_p)^{2N}} \quad (1)$$

Calculating the arithmetic mean of the signal samples, the DC level was eliminated.

Next, at the third block, for each V6 derivation signal, the signal energy was calculated according to Equation 2.

$$\sum_{n=n_1}^{n_2} |x[n]|^2 \quad (2)$$

Were |x| denotes the magnitude of the (possible complex) number x (Oppenheim;

Willsky; NAWAB, 1996). Given that the electrocardiogram signal is periodic, this calculated energy is finite.

In this block, the signal transform was also calculated and the frequency with the largest spectrum of the signal was identified. In this way, the feature pairs of all V6 derivation signals were stored for validation of the neural network.

In the fourth block, a neural network with fifty neurons was implemented. Each neuron has two inputs and one output. The two network inputs were fed with pairs of the features extracted from each signal. The neural network in this block uses structures with only one hidden layer with fixed and possibly random weights, which allows for representing complex non-linear functions (universal approximators). This type of architecture has the convenience of only needing to adjust the weights of the output layer, which in addition to reducing the memory also facilitates the learning process that can be carried out only with a multiple linear regression algorithm. Therefore, the desired target function can be obtained avoiding the computationally intensive process of backpropagation during the learning phase. This type of network Neural is also known as Extreme Learning Machine. Initially, the network was trained with the application of 120 signals with 68 normal diagnoses and 52 with a heart attack diagnosis. After training the neural network, 480 signals were applied for verification, 161 with a normal diagnosis and 319 with a diagnosis of infarction.

In the final block, all output values from the neural network were collected, which correspond to the diagnoses for system validation. It was standardized such that for a normal electrocardiogram diagnosis, the network should generate a value of zero and for a heart attack diagnosis, the network should generate a value of 1. With this information, graphs, and calculations were generated to quantify the efficiency of the system under study.

## Results and Discussion

Figure 3 shows the 10000 samples from the V6 lead of an electrocardiogram signal taken from the PhysioNet database (PhysioNet, 2023).

Figure 3. Samples of an electrocardiogram signal.

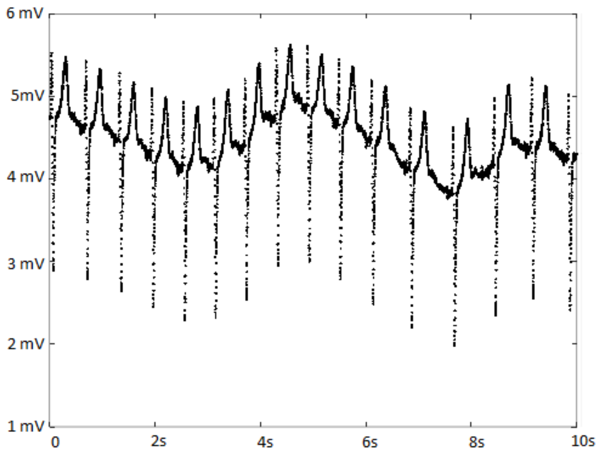
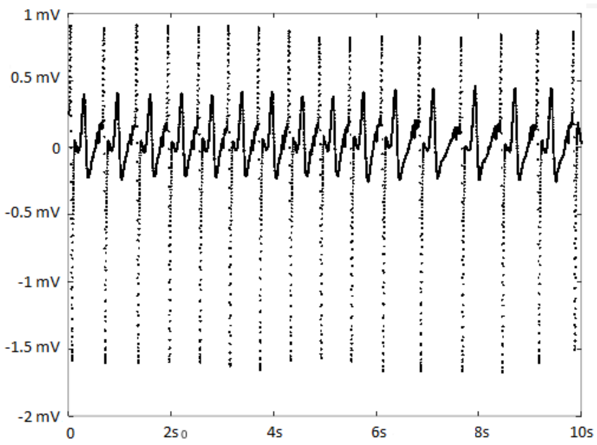



Figure 4 shows the 10000 samples after the signal filtering process.

Based on the samples depicted in Figure 4, the average energy of this signal was calculated (using Equation 2).

Figure 4. Samples of an electrocardiogram after filter.




The other input applied to the neural network architecture is the strongest frequency in the spectrum of the evaluated electrocardiogram signal.

As an example, for the electrocardiogram signal in Figure 4, its transform was calculated and the frequency with the largest spectrum was selected. The highest intensity spectrum frequency for the signal under analysis was 33Hz.

That way, for the electrocardiogram signal being analyzed in this example, the values of 1,5852.104 (signal energy) and 33Hz (main frequency) would be applied as inputs into the neural network.

To validate the presented method the system validation received 68 normal electrocardiogram signals and 52 infarction signals for network training.

After the neural network's training,161 signals from the V6 lead with normal diagnoses and 319 signals from the V6 lead with an infarction diagnosis were applied.

The PhysioNet offers free access to several databases on physiological signals with medical opinions and allows simulations and tests with real ECG signals (PhysioNet, 2023).

According to the tests carried out in the simulation, Figure 5 shows the output surface generated by the network after training, as a function of the two inputs. To facilitate the network's implementation, the input values were normalized to a range between o and the network's outputs in the training sets were normalized to have a normal electrocardiogram represented by 5 and infarction by 15.

Figure 1. Surface generated from network training.

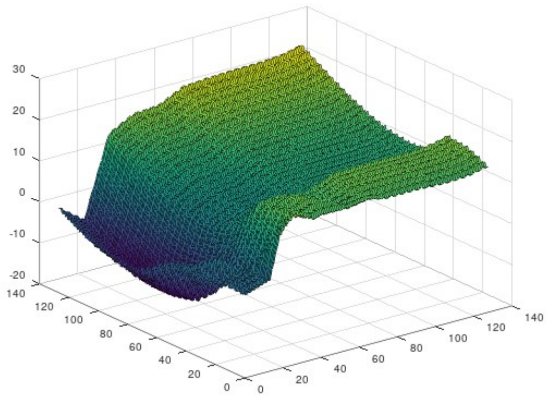


Figure 6 shows the positions and values of the network training signal outputs.

Figure 2 – Expected output of signals used for network training.

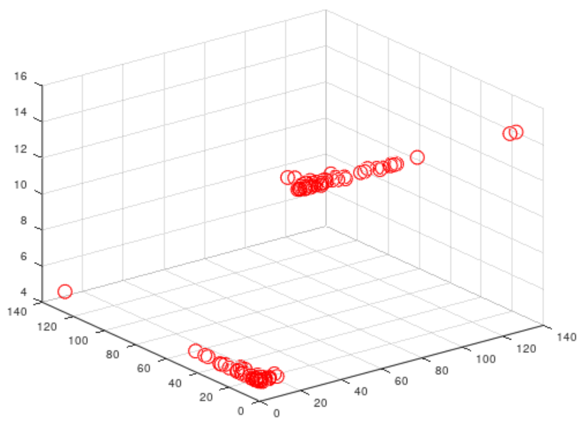


Figure 7 shows the network output of the new test signals used to validate the operation of the network.

Based on the obtained results, an evaluation of the network's performance was conducted. Figure 8 presents the errors generated in comparison to the expected correct results.

Figure 3. Network output of signals used for network validation.

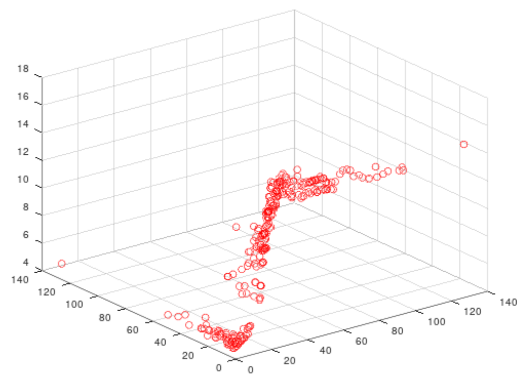
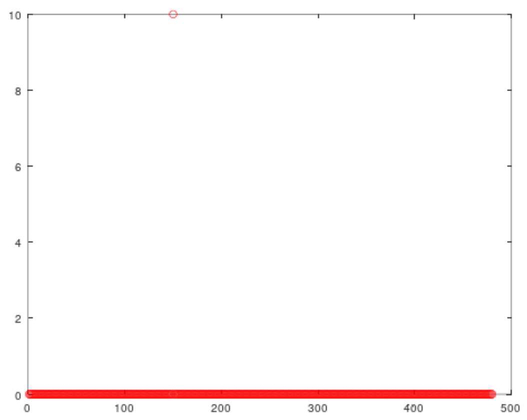


Figure 4. Errors generated at Network output.



According to the error graph presented, it was found that of the 480 electrocardiogram signals tested, 479 were correctly diagnosed by the network.

To validate the result obtained with other works in the literature, figures of merit were calculated in terms of accuracy (Acc). The Acc parameter indicates the accuracy of the system under test and can be obtained according to Equation 3 (Inan; Giovangrndi; Kovacs, 2006).

$$Acc = 1 - N_{err}/N_{total} \qquad (3)$$

Where $N_{err}$ indicates the number of misdiagnoses and $N_{total}$ indicates the total number of diagnoses.

The Se parameter indicates the percentage of correct diagnoses compared to undetected issues. It can be obtained according to the Equation 4 (Inan Omer T., et al, 2006).

$$Se = Tp/(Tp + Fn) \qquad (4)$$

Where Tp and Fn are the number of correct and undetected diagnoses, respectively.

The Pp parameter indicates the percentage of correct diagnoses in comparison with incorrect diagnoses. It is given by Equation 5 (Inan Omer T., et al, 2006).

$$Pp = Tp/(Tp + Fp) \qquad (5)$$

Where Fp indicates the number of incorrect diagnoses.

Table 1 compares previous works and their respective methods with the results obtained using the proposed method for diagnosing electrocardiograms.

Based on the validation system developed and the parameters to be compared, tests were performed to verify how effective the tool is. A comparison with other implemented systems in the literature is also show in table 1.

Table 1. Comparison of previous jobs with current job

| | Channels | Accuracy | Pp | Se |
|---|---|---|---|---|
| Gupta | 1 | - | - | 99,92%- |
| Maglaveras | 4 | - | 78% | 88% |
| Jager | 2 | - | 87.1% | 83.8% |
| Anuradha | 1 | 93.13% | - | - |
| Zheng | 1 | - | 99.59% | 99.68% |
| Sharma | 1 | 99% at 100% | - | - |
| Chen | 1 | 99.65% | - | - |
| Afsar | 1 | 96.75% | - | - |
| This Work | 1 | 99.79% | 100% | 99,73% |

The results demonstrate that the proposed system achieved comparable or superior performance across all parameters (Se, Pp, Acc). Notably, this performance was obtained using only a single lead from the patient's electrocardiogram, and the processing required to extract input features for the network is straightforward. This simplifies both the implementation and usage of a potential hardware solution. In this case, only one electrode connection to the patient is required, corresponding to the V6 lead.

## Conclusions

The proposed system demonstrates the capability to offer probable diagnoses with comparable effectiveness to existing processing systems, achieved through the utilization of neural networks. By applying this neural network framework, we could extract relevant features from a singular derivation of an electrocardiogram, enabling the generation of diagnoses with similar or better performance then those reported in previous studies. According to the calculations detailed in this work, the system can effectively diagnose infarction using just a single derivation and simple processing of an electrocardiogram signal. This way it is possible to implement simpler hardware with less processing.

## Acknowledgment

This work has been supported by the following Brazilian research agencies: CAPES, CNPq, and FAPEMIG.

## References

AFSAR, F. A.; AKRAM, M. U.; ARIF, M.; KHURSHID, J. A Pruned Fuzzy k-Nearest Neighbor Classifier with Application to Electrocardiogram Based Cardiac Arrhythmia Recognition. Proceedings of the 12th IEEE International Multitopic Conference, December 23-24, 2008.

ALKHODARI, M.; FRAIWAN, L. Convolutional and recurrent neural networks for the detection of valvular heart diseases in phonocardiogram recordings. ELSEVIER Computer Methods and Programs in Biomedicine. 2021.

ANURADHA, B.; REDDY, V. C. V. Cardiac Arrhythmia Classification Using Fuzzy Classifiers. Journal of Theoretical and Applied Information Technology. 2005 – 2008.

CHEN, Y. H.; YU, S. N. Comparison of Different Wavelet Subband Features in the Classification of ECG Beats Using Probabilistic Neural Network. EMBS Annual International Conference New York City, USA, Aug 30-Sept 3, 2006.

DALLY, W. J.; GRAY, C. T.; PAULTON, J.; KHAILANY, B.; WILSON, J.; DENNISON, L. Hardware-Enable Artificial Intelligence. IEEE Symposium on VLSI Circuits. 18-22 June 2018.

DI CESARE M.; BIXBY H., GAZIANO T.; HADEED L.; KABUDULA C.; MCGHIE D. V.; MWANGI J.; PERVAN B.; PEREL P.; PINEIRO D.; TAYLOR S.; PNTO F. World heart report 2023: Confronting the world's number one killer. World Heart Federation, Geneva, Switzerland, 2023.

GUPTA C.; MITTAL M.; MITTAL V. Chaos Theory: Na Emerging Tool for Arrhythmia Detection. Sensing and Imaging, 2020.

GUPTA V.; MITTAL M.; MITTAL V.; GUPTA A. An eficiente AR modelling-based electrocardiogram signal analysis for health informatics. Int. J. Medical Engineering and Informatics. Vol. 14, No. 1, 2022.

HAMPTON J.; HAMPTON J. The ECG Made Easy. Ninth edition. Elsevier, 2019.

OPPENHEIM A. V.; WILLSKY A. S.; NAWAB W. S. H. Signals & Systems. Second Edition, Prentice Hass, August 1996.

INAN, O. T.; GIOVANGRNDI, L.; KOVACS, G. T. A. Robust Neural-Network-Based Classification of Premature Ventricular Contractions Using Wavelet Transform and Timing Interval Features. IEEE Transactions on Biomedical Engineering. V. 53, n. 12, 2006.

JAGER, F.; MOODY G. B.; TADDEI A.; and R. G. MARK G. Performance measures for algorithms to detect transient ischemic ST segment changes.In Computers in Cardiology. Los Alamitos, CA: IEEE Comput. Soc. Press. P. 369–372, 1991.

JAGER F.; MOODY G. B.; MARK R. G. Detection of transient ST segment episodes during ambulatory ECG monitoring. Comput Biomed Res. 31(5):305-22, 1998.

MAGLAVERAS, N.; STAMKOPOULOS T. et. Al. An adaptive backpropagation neural network for real-time ischemia episodes detection: development and performance analysis using the European ST-T database. IEEE Trans Biomed Eng.45(7):193-214, 1998.

OPPENHEIM, A. V.; WILLSKY, A. S., NAWAB, W. S. H. Signals & Systems. Prentice Hass. 2. Ed., 1996.

PATEL I.; SANDHYA A.; RAJA, V. S.; SARAVANAN S. Extraction of Features from ECG Signal. International Journal of Current Research and Review. Vol. 13, 2021.

Physionet PhysioBank MIT-BIH Arrhythmia Database. Disponível em: <www.physionet.org/physiobank/database/mitdb>. Acessado em 04 de setembro de 2023.

RAHMAN, S.; PAL, S.; YEARWOOD, J.; KARMACAR, C. Robustness of Deep Learning models in electrocardiogram noise detection an classification. ELSEVIER Computer Methods and Programs in Biomedicine. 2024.

SEDRA, Adel S., SMITH, Kenneth, C., Microeletrônica. 4. Ed. – São Paulo: Makron Books, 2000.

SHARMA, L. D.; SUNKARIA, R. K. A robust QRS detection using novel pre-processing techniques and kurtosis based Enhanced dfficiency. Measurement – ELSEVIER. 2016.

SHEN, S.; CHEN, J.; WANG, W.; TSAI, D.; SHEN, L.; TSENG, C. A CNN-Based Human Head Detection Algorithm Implemented on Edge AI Chip. IEEE, International Conference on System Science and Engineering (ICSSE). 31 August 2020 – 03 September 2020.

ZHENG, H.; WU, J. Real-time QRS Detection Method. 10th IEEE Intl. Conf. on e-Health Networking. Applications and Service (HEALTHCOM 2008).