Cardiac Arrhythmia Detection Based on R-Peak Centered Segments of ECG Signals using 1D Convolutional Neural Networks and Explanation using Grad-CAM Tool

A. S. do Canto[1], F. A. P. Januário[1], M. G. F. Costa[1], C. F. F. Costa Filho[1]

[1] Center for R&D in Electronics and Information/Federal University of Amazonas, Manaus, Brazil

Abstract

This work proposes a method that combines machine learning and signal processing techniques to detect and classify cardiac arrhythmias in ECG signals, using the MIT-BIH Arrhythmia dataset. The methodology consists of the following steps: preprocessing to eliminate atypical morphologies in MLII derivation or accelerated heart rhythm, segmentation of the R-Peak annotations, and arrhythmias classification using 1D CNN (10 different types of arrhythmias were classified). The model achieved the following performance: accuracy and precision - 99.40%, recall - 99.32%. For explaining the results, the Grad-CAM technique was used for interpretability, identifying relevant ECG regions in the decisions. The results prove the effectiveness of the method in the automatic detection of arrhythmias.

Keywords — machine learning, arrhythmias, 1D CNN, R-Peak, Grad-Cam.

1. Introduction

Cardiovascular diseases (CVDs) are the leading cause of death worldwide, accounting for 17.9 million deaths in 2019, and for 32% of global deaths. Most cases occur in low- and middle-income countries, and 85% of deaths are related to heart attacks and strokes. Many of these deaths can be prevented by reducing risk factors such as smoking, unhealthy diet, and physical inactivity. Early detection and appropriate treatment are essential to reduce CVD mortality [1].

Rajpurkar et al. [2] report that manual analysis of these signals is complex and error-prone, especially in continuous monitoring. Automated arrhythmia detection systems have emerged as a promising approach to improve diagnostic accuracy and reduce response time.

Acharya et al. [3] developed a 9-layer deep convolutional neural network (CNN) to automatically classify five categories of cardiac segments from 260 samples in ECG signals. The model was trained on raw data processed to remove noise, as well as an artificially augmented set to balance the classes. In the testing, CNN achieved an accuracy of 94.03% on raw ECGs and 93.47% on noise-free signals, while with unbalanced data, the accuracy dropped to approximately 89%. These results demonstrate the potential of CNN as an auxiliary tool in the detection of automatic arrhythmias.

Zhou et al. [4] proposed a hybrid method that combines convolutional neural networks (CNN) with extreme learning machines (ELM) for the automatic classification of four classes of arrhythmias in ECG signals. The approach aims to mitigate the challenges caused by noise and poor signal quality, improving accuracy in the diagnosis of arrhythmias. The methodology includes segmenting the signal around the R peak of the QRS wave with 250 samples, ensuring a more accurate analysis of the heartbeat. The experiments demonstrated that the model achieved an accuracy rate of 98.77%, which demonstrates high generalizability for different data sets.

Ahmed et al. [5] proposed a deep learning architecture based on one-dimensional convolutional neural networks (1D-CNN) for the automatic classification of four types of cardiac arrhythmias. The model was trained with real signals from the MIT-BIH database, previously processed for noise reduction. The methodology involved the extraction of beats from ECG lead II, using normalization and segmentation techniques in 180-sample windows based on R-peak detection. The results showed satisfactory performance, with 100% accuracy in training and 99.0% in testing, standing out as an efficient alternative for the automated diagnosis of arrhythmias.

This study presents a robust model for automatic classification of heartbeat-associated arrhythmias, focusing on the analysis of P, QRS, and T waves. It combines signal preprocessing, R-peak-based segmentation, and a deep 1D-CNN architecture. The method adapts the segmentation to specific morphologies of arrhythmias, increasing the sensitivity to the associated arrhythmia, thus providing a possible scalable clinical solution.

2. Methodology

2.1. Database Processing

The MIT-BIH Arrhythmia Dataset [6] was used. This dataset contains 48 ECG recordings (30 minutes each, sample rate of 360 Hz) from 47 patients, with manual annotations of arrhythmias by specialists. This dataset was chosen for its reliability, diversity and recognition as a benchmark in signal processing and machine learning. In the preprocessing phase, illustrated in Figure 1, we identified, for each record, the leads present in the MIT-BIH Arrhythmia dataset, selecting exclusively records with MLII, the gold standard for arrhythmia analysis. The last block

Figure 1. Steps of data preprocessing

in Figure 1, aims to eliminate records with atypical morphologies or accelerated cardiac rhythms. Atypical morphologies were observed in classes NB, LBBB, RBBB and PAC, even within the same derivation(MLII). Another atypical condition identified was the presence of short RR intervals and multiform PVCs. Figure 2 illustrates these problems in records 111, 108, 124, and especially in 207, which has the highest number of atypical conditions between LBBB, RBBB, PAC and PVC classes. Figure 2(a)

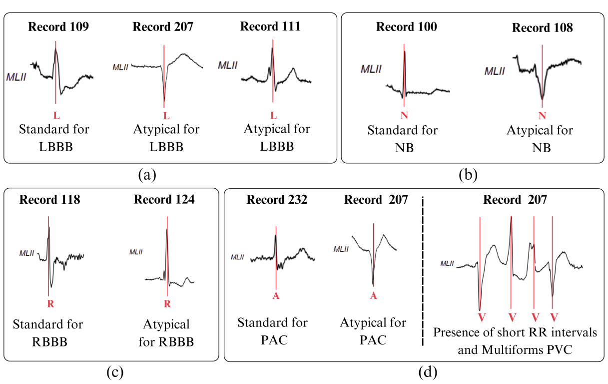

Figure 2. Classes with atypical morphologies (a) LBBB, (b) NB, (c) RBBB and (d) PAC and PVC. In Figure (d) the multiple PVC morphologies and short RR intervals are evidenced in record 207.

presents a standard LBBB class(MLII, record 109) and atypical morphology (records 207 and 111). Figure 2(b) shows a standard NB class(MLII, record 100) with atypical morphology (record 108), while Figure 2(c) shows a standard RBBB class (MLII, record 118) and atypical morphology (record 124). Figure 2(d) presents a standard PAC class(MLII, record 232) with atypical morphology in record 207, which also presents short RRs and multiple PVC morphologies.

2.2. Segmentation and Normalization

The use of preexisting annotations from the MIT-BIH arrhythmia dataset, which provide the exact positions of the R peaks, significantly simplifying the segmentation process. According to Malmivuo [7], the complete duration of cardiac events lasts 600 ms. Based on Malmivuo description, the ECG signals were divided into 600 ms segments (216 samples corresponding to a sampling rate of 360 Hz), centered on the R peaks, with asymmetric windows of 212.5 ms (76.5 samples) before the R peak and 387.5 ms (139.5 samples) after the R peak, ensuring the capture of the morphological characteristics of the P, QRS and T waves. As illustrated in Figure 3, this approach ensures the faithful representation of the rhythmic and morphological behavior of the signal, while maintaining the temporal and physiological consistency required for arrhythmia analysis.

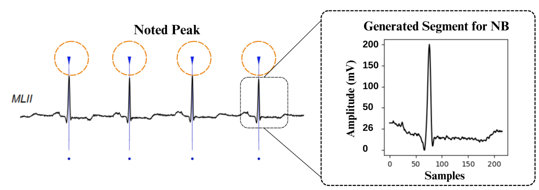

Figure 3. Unnormalized Generated Segment

As demonstrated by [8], data normalization is essential in pattern recognition, whether supervised or not supervised learning is employed. This work used minmax normalization to reduce amplitude variations, preserving morphological characteristics, as illustrated in the process in Figure 4. The data set was partitioned into training subsets (70%),

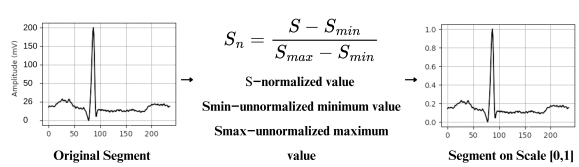

Figure 4. Data Normalization

validation subsets (15%), and test subsets (15%) while preserving the original proportion of arrhythmia classes to prevent bias and ensure the generalizability of the model. The detected arrhythmias, a as illustrated in Figure 5, include Normal Beat (NB, N), Premature Atrial Contraction (PAC, A), Fusion of Ventricular and Normal Beat (FVNB, F), Fusion of Paced and Normal Beat (FPNB, f), Left Bundle Branch Block (LBBB, L), Right Bundle Branch Block (RBBB, R), Premature Ventricular Contraction (PVC, V), Paced Beat (PB, /), Aberrated Atrial Premature (AAP, a), and Nodal Escape Beat (NEB, j).

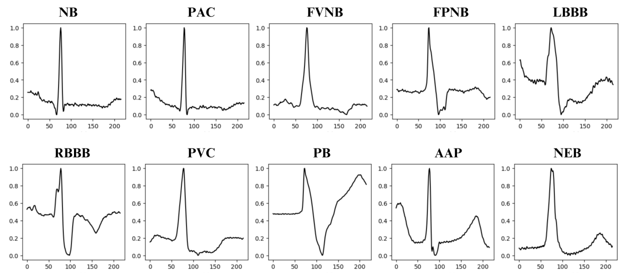

Figure 5. Generated Segments

2.3. Neural Network Architecture: Training and Testing

The proposed neural network, illustrated in Figure 6, employs a sequential architecture with three one-dimensional convolutional blocks (96, 128, and 256 filters, kernel=10, stride=1, padding='same') interleaved with batch normalization and ReLU, followed by max pooling (window=5); L_2 regularization is applied in the last three convolutional layers. A flatten layer is used for vectorization, followed by two dense layers (128/96 neurons, ReLU + L_2), and a final softmax layer for multiclass classification. The model was trained using the Adam optimizer

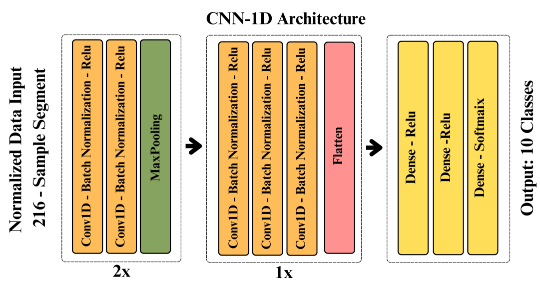

Figure 6. Proposed Architecture

(initial learning rate = 1e-3) with the categorical cross-entropy loss function, monitoring the precision, recall and accuracy metrics for multiclass evaluation. Three essential callbacks were implemented: Early Stopping (patience = 20 epochs) to prevent overfitting. Learning Rate Scheduler for dynamic adjustment of the learning rate. Model Checkpoint to save the best model based on the validation loss. The training employed batches of 512 samples (with shuffle), class weights for balancing, and lasted a maximum of 100 epochs, using an independent validation set to ensure model generalization.

3. Results and Discussion

Table 1 shows the results obtained for the classification of 10 classes. The following metrics were obtained: accuracy, recall,and precision. The confusion matrix, shown in Table 2, evaluates the performance of the model when comparing the true labels with the predictions made. The principal diagonal values represent the correct predictions for each class, while the secondary diagonal values indicate misclassification.

Table 1. Comparison of results with works using R-peak-based segmentation and CNN-1D

Reference	Method	Acc.(%)	Recall (%)	Prec. (%)
Proposed Model	1D CNN, 10 classes and 216 samples	**99.40**	**99.32**	**99.40**
Acharya et al. [3]	1D CNN, 5 classes and 260 samples	94.03	96.71	-
Zhou et al. [4]	1D CNN + ELM, 4 classes and 250 samples	98.77	-	-
Ahmed et. al [5]	1D CNN, 4 classes and 180 samples	99.00	94.00	-

Table 2. Confusion Matrix with Absolute Values

True Label	Predicted Label									
	N	A	F	f	L	R	V	/	a	j
N	10671	17	3	0	0	0	7	0	0	2
A	25	335	0	0	0	1	0	0	0	0
F	9	0	98	0	0	0	13	0	0	0
f	0	0	0	39	0	0	0	0	0	0
L	2	0	0	0	673	0	0	0	0	0
R	0	1	0	0	0	847	0	0	0	0
V	17	0	5	0	0	0	898	0	0	0
/	0	0	0	0	0	0	0	543	0	0
a	4	0	0	0	0	2	1	0	18	0
j	10	1	0	0	0	0	0	0	0	24

3.1. Grad-Cam Explanation

Gradient-weighted Class Activation Mapping (Grad-CAM) was used to explain model decisions by highlighting signal regions relevant to arrhythmia classification. Using gradients from the last CNN layer [9], it enables evaluation of the clinical relevance of detected patterns. Figure 7 presents Grad-CAM results for LBBB and NEB. In

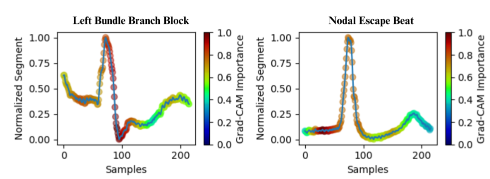

Figure 7. Grad-Cam for LBBB and NEB

LBBB, activation focuses on the wide QRS and altered ST-T interval, confirming effective model learning. For NEB, attention appears in the irregular QRS and short or absent PR interval, aiding distinction from other arrhythmias. Figure 8 shows the Grad-CAM results for NB, PAC, RBBB, and PVC. In normal beats, activation is centered in the QRS, with balanced attention in ST and T waves, confirming sinus rhythm recognition. In PAC, focus is on the premature P wave, its key marker. RBBB shows activation in the widened QRS and altered repolarization. PVC displays strong activation in the premature, widened QRS and asymmetric repolarization, confirming correct classification. Figure 9 presents the Grad-CAM analysis for

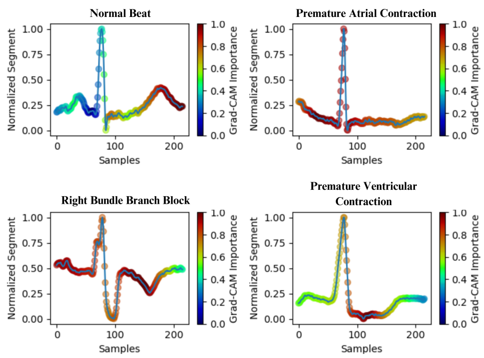

Figure 8. Grad-Cam for NB, PAC, RBBB, and PVC

FNVB, FPNB, PB, and AAP. For FNVB, activation appears in QRS regions distinct from each isolated beat, emphasizing transition zones and confirming correct pattern recognition. For FPNB, intense coloring highlights pre-QRS and post-QRS regions, indicating the model's strong performance. In PB, activation focuses on the stimulation spike, widened QRS, and repolarization changes, confirming artificial pacing. In APP, the predominant attention is on the premature P wave and the wide QRS complex, demonstrating the good generalization of the model.

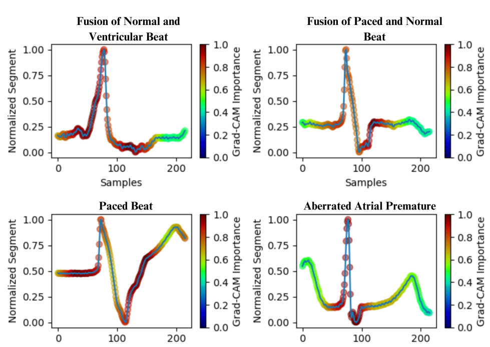

Figure 9. Grad-Cam for FNVB, FPNB, PB, AAB

4. Conclusion

The proposed method exceeded the state of the art, with an accuracy of 99.40% in the detection of arrhythmias. The analysis using Grad-Cam has effectively demonstrated the model capacity to accurately focus on the heartbeat regions that have relevant standards associated with each class anomalies. The results indicate that the model was able to properly generalize the patterns related to the main characteristics of a heartbeat (P complex, QRS and T wave), both in the identification of normal beats and other arrhythmias.

References

[1] World Health Organization. (n.d.). *Cardiovascular diseases (CVDs)*. World Health Organization. Retrieved April 1, 2025, from `https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)`

[2] P. Rajpurkar, A. Y. Hannun, M. Haghpanahi, C. Bourn, and A. Y. Ng, Cardiologist-Level Arrhythmia Detection with Convolutional Neural Networks, 2017. DOI: `https://doi.org/10.48550/arXiv.1707.01836`

[3] Acharya, U. Rajendra et al. *A deep convolutional neural network model to classify heartbeats*. Computers in Biology and Medicine, v. 89, p. 389-396, 2017. ISSN 0010-4825. DOI: `https://doi.org/10.1016/j.compbiomed.2017.08.022`.

[4] Zhou, Shuren; TAN, Bo. *Electrocardiogram soft computing using hybrid deep learning CNN-ELM*. Applied Soft Computing, v. 86, p. 105778, 2020. ISSN 1568-4946. DOI: `https://doi.org/10.1016/j.asoc.2019.105778`.

[5] Ahmed, Adel A.; ALI, Waleed; ABDULLAH, Talal A.; MALEBARY, Sharaf J. *Classifying Cardiac Arrhythmia from ECG Signal Using 1D CNN Deep Learning Model*. Mathematics, v. 11, n. 3, p. 562, 2023. ISSN 2227-7390. DOI: `https://doi.org/10.3390/math11030562`.

[6] Moody, George B and Mark, Roger G. *The impact of the MIT-BIH arrhythmia database*. IEEE engineering in medicine and biology magazine. 2001;20:45–50. DOI: `https://doi.org/10.1109/51.932724`

[7] MALMIVUO, Jaakko; PLONSEY, Robert. *Bioelectromagnetism: principles and applications of bioelectric and biomagnetic fields*. New York: Oxford University Press, 1995.

[8] L. da Fontoura Costa, Data Normalization in Signal and Pattern Analysis and Recognition: A Modeling Approach, working paper, Jun. 2022. [Online]. Available: `https://hal.science/hal-03688208`

[9] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization, *International Journal of Computer Vision*, vol. 128, no. 2, pp. 336–359, Oct. 2019. DOI: 10.1007/s11263-019-01228-7.

Address for correspondence:

Aris Souza Do Canto
ariscanto21@gmail.com