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 to detect and classify cardiac arrhythmias in ECG signals, using the MIT-BIH Arrhythmia dataset. The methodology includes preprocessing to eliminate atypical morphologies and accelerated rhythms, inconsistencies often overlooked in previous studies that can lead to feature overlap and compromise model accuracy, asymmetric 600 ms segmentation centered on the R peak to capture the features of the P, QRS, and T waves, and classification of 10 types of arrhythmias, covering a broader and more challenging scenario compared to studies that usually consider only 4 to 5 classes. The model achieved high performance, with 99.40% accuracy and precision and 99.32% recall. The Grad-CAM technique was applied to confirm that the model focuses on clinically relevant regions of the ECG, increasing interpretability and clinical confidence.

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 1D-CNN architecture for the automatic classification of four types of cardiac arrhythmias, using signals from the MIT-BIH database previously processed for noise reduction. The methodology involved the extraction of beats from ECG lead II, applying normalization and segmentation in 180-sample windows centered on R-peak detection. The model achieved excellent performance, with 100% accuracy in training and 99.0% accuracy in testing, standing out as an efficient and promising alternative for automated arrhythmia diagnosis. This study proposes a robust model for automatic classification of cardiac arrhythmias, combining pre-processing, non-peak centered asymmetric segmentation R and deep 1D-CNN. The model adjusts segmentation to capture specific morphologies and uses Grad-CAM to visually highlight relevant ECG regions, increasing interpretability and clinical confidence.

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.

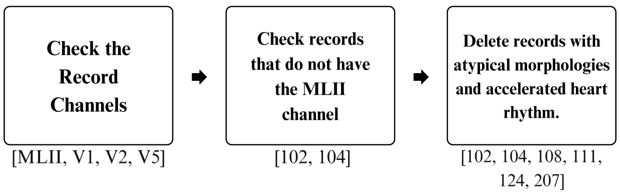

Figure 1: Steps of data preprocessing.

The last block in Figure 2, 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.

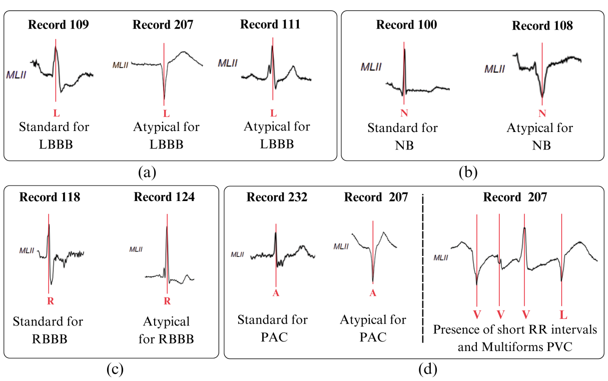

Figure 2: Classes with atypical morphologies (a) LBBB, (b) NB, (c) RBBB and (d) PAC and PVC. In Figure (d) the multiple PVC morphologies and short RR intervals are evidenced in record 207.

Figure 2(a) 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 pre-existing annotations from the MIT-BIH arrhythmia dataset, which provide the exact positions of the R peaks, significantly simplifies the segmentation process. According to Malmivuo [7], the full duration of cardiac events is 600 ms. Based on this 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. This asymmetric window, which differs from previous studies that used shorter or longer windows, was designed to ensure the complete capture of the morphological features of the P wave, QRS complex, and T wave. As illustrated in Figure 3, this approach provides a physiologically consistent representation of the signal, allowing for more accurate arrhythmia classification.

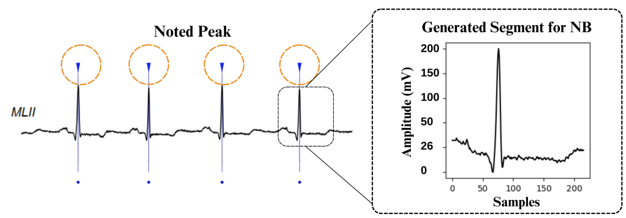

Figure 3: Unnormalized generated segment.

As demonstrated by [8], data normalization is essential in pattern recognition, whether supervised or unsupervised learning is employed. This work used min-max normalization to reduce amplitude variations, preserving morphological characteristics, as illustrated in the process in Figure 4. To ensure model robustness and generalizability, the dataset was stratified and partitioned into training (70%), validation (15%), and testing (15%) subsets. Importantly, this partitioning was subjective, meaning that all segments belonging to a specific arrhythmia class from a single patient record were exclusively allocated to a subset. This careful separation strategy preserved the original class distribution, avoiding data leakage and overestimation of results, common problems when segments from the same patient

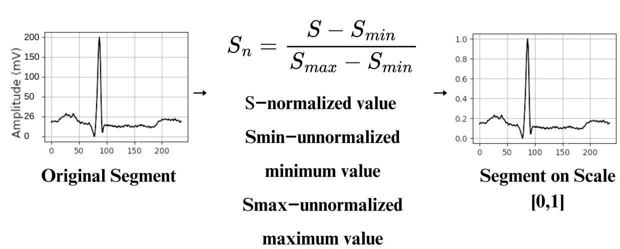

Figure 4: Data normalization.

appear in both the training and testing sets. The detected arrhythmias, 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).

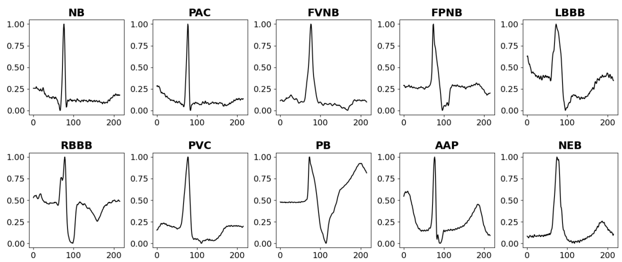

Figure 5: Example of the generated segments.

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.

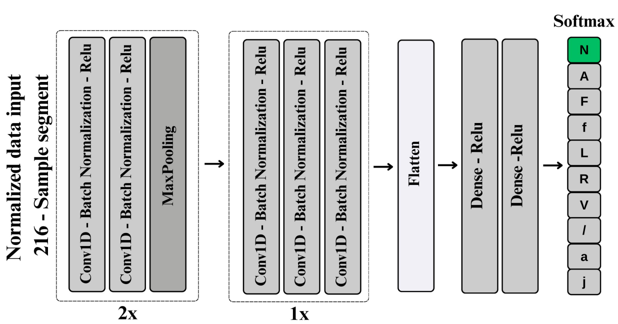

Figure 6: Proposed architecture.

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 (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.

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	-

The confusion matrix, shown in Figure 7, 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.

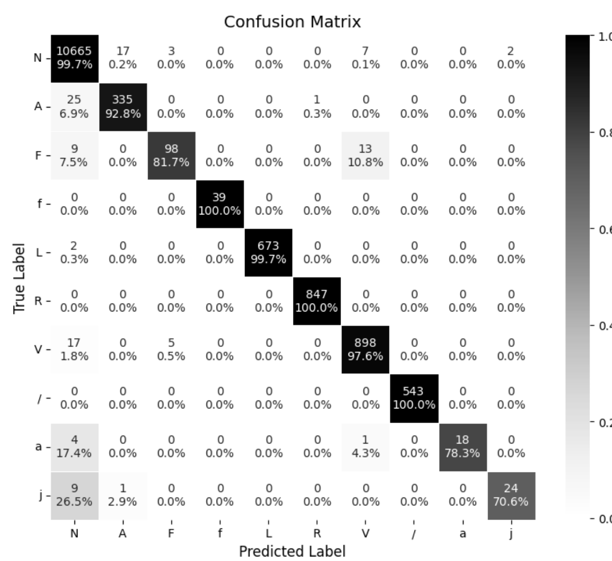

Figure 7: Confusion matrix with absolute values.

3.1. Grad-Cam Explanation

To enhance interpretability and clinical confidence in the model's decisions, the Grad-CAM (Gradient-weighted

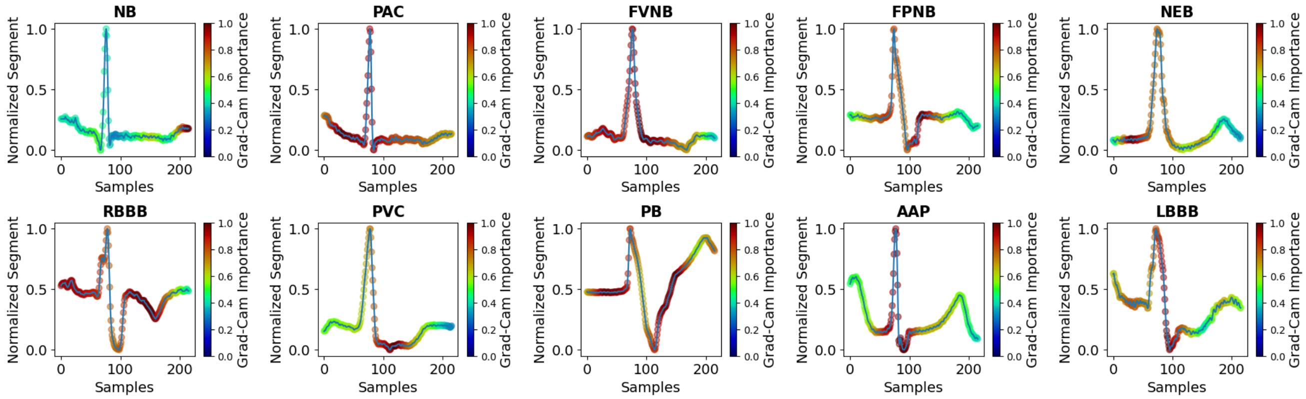

Figure 8: Grad-Cam for NB, PAC, RBBB, PVC, FNVB, FPNB, PB, AAP, NEB and LBBB

Class Activation Mapping) technique was applied [9]. Complementing the numerical performance metrics, Grad-CAM allows for an in-depth visual analysis, identifying the ECG signal regions that contributed most to the classification of each segment. Figure 8 presents the activation maps for the 10 classes, confirming that the model consistently focuses on clinically significant ECG regions and has learned to identify relevant morphological markers. The observed activation regions were: on the normal QRS for NB; on the premature P wave for PAC; on the wide QRS and altered repolarization for RBBB; and on the premature, wide QRS with asymmetric repolarization for PVC. Furthermore, for FNVB and FPNB, activation occurred in distinct QRS regions and with intense pre- and post-QRS highlighting, respectively. For PB, the focus occurred on the pacemaker spike and the subsequent wide QRS, while for AAP, activation concentrated on the premature P wave followed by a wide QRS. Finally, for LBBB and NEB, the focus was on the wide QRS with ST-T changes and on the escape QRS with a short or absent PR interval, respectively.

4. Conclusion

The proposed method outperformed the state-of-the-art, achieving 99.40% accuracy in arrhythmia detection. Grad-CAM analysis confirmed the model's ability to precisely target regions of the heartbeat with class-specific patterns. The results show that the model effectively generalized the main features of the heartbeat (P wave, QRS complex, and T wave), correctly identifying normal and arrhythmic beats. Grad-CAM consistently aligned with clinical expectations, improving interpretability and reinforcing its potential as a reliable diagnostic tool.

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.

[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.

[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.

[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.

[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.

[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. 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.

Address for correspondence:

Áris Souza Do Canto
Center for R&D in Electronics and Information
Manaus, AM, 69067005, Brazil
ariscanto21@gmail.com