Real-time Heartbeat Classification Based on Parallel QRS Detection

Wojciech Reklewski[1], Jeremiasz Potoczny[1], Marek Miśkowicz[1], Piotr Augustyniak[1]

[1]AGH University of Krakow, Krakow, Poland

Abstract

The main challenge for automatic classification of cardiac arrhythmias with battery-operated devices is to achieve good classification results without high computational load and related high energy consumption. In this paper, we present a heartbeat classification method that uses the differences in R-peak time locations detected by multiple low power QRS detection algorithms operating in parallel. The outputs of the detectors are the inputs to the decision tree classifier. The classification is divided into three QRS morphology types: (1) Normal and Atrial Premature, (2) Ventricular, and (3) Other. The overall accuracy (ACC) of the proposed classification method for test dataset is 90.52%. The detailed results are as follows: Positive Predictivity (PPV) 93.82%, Sensitivity (Se) 95.51% for Normal and Atrial Premature class; PPV=84.66%, Se=76.33% for Ventricular class; PPV=84.33%, Se=79.29% for Other class. The proposed arrhythmia classification method is applicable for real-time mobile ischemia detection and HRV analysis.

1. Introduction

Wearable electronic devices are becoming increasingly comparable to professional health care devices [1]. While the primary design objective for professional medical equipment is to maximize performance, for wearable battery-operated devices, the main concerns are power consumption, available memory, device size and weight. These factors should be optimized possibly without compromising diagnostic performance. Another design criterion is the type of output received from the device, whether it is raw data, preprocessed data, or disease detection supporting medical diagnosis. The continuous monitoring of patients' heart during their daily activity began with the invention of the Holter ECG monitors in 1949 [2]. Nevertheless, standard Holter ECG devices produce large volumes of data, necessitating significant amounts of a doctor's time for data analysis before reaching a diagnosis. The research on automatic detection of arrhythmia and other cardiac disorders accelerated with the computer revolution of the 1970s and continues to progress with machine learning and data mining [3, 4].

The automatic ECG-based arrhythmia detection classification consists of four processing steps: (1) ECG signal preprocessing, (2) QRS detection and signal heartbeat segmentation, (3) feature extraction and (4) arrhythmia classification [3].

The ECG signal preprocessing and QRS detection are subject of research since decades and many algorithms have been proposed for ambulatory equipment [5] with battery-operated devices [6].

Feature extraction is the key stage in automatic arrhythmia classification [3]. It involves extracting information from heartbeats that can be used for classification. The most popular features include heartbeat interval, QRS complex duration, and points of the segmented ECG curve. Other methods include linear predictive coding, high-order accumulates, clustering, correlation dimension and largest Lyapunov exponent, Hermite transform, and local fractal dimension. The best performance results are reported using wavelet transform methods [3], although these require computationally intensive calculations.

Approaches to arrhythmia classification can be grouped into five main categories: support vector machine (SVN), artificial neural network (ANN), linear discriminant (LD), Reservoir Computing with Logistic Regression (RC), and other approaches (e.g., decision tree, nearest neighbor, clustering, hidden Markov models, hyperbox classification, optimum-path forest, conditional random fields and rules-based models) [3].

Despite significant research efforts, automatic classification of cardiac arrhythmias remains an open research problem due to high computational cost, difficulties related to machine learning with unbalanced databases, lack of databases with sufficient size and diversity, and the absence of standard evaluation protocols as evaluation schemes impact the performance results [3].

The objective of this study was to develop an effective algorithm for automatic arrhythmia detection with low computational cost for battery-operated ECG devices.

2. Materials and Methods

The data from MIT-BIH Arrhythmia Database (MIT-BIH AD) [7] was used in this work. The database

 ISSN: 2325-887X DOI: 10.22489/CinC.2024.156

consists of 48 half-hour recordings of ECG signal, with two channels. Upper channel data that is lead II was used for this work. The database consists of close to 110 000 heartbeats that was enough for machine learning approach.

Algorithm classification results were visualized in confusion matrix with three categories and analyzed with following parameters: overall *Accuracy* (*ACC* = *Number of Correct Predictions* / *Total Number of Predictions*), and after reduction of multiclass classification to binary classification by converting the multiclass problem to one class vs many classes by standard binary classification parameters: *Positive Predictive Value* (*PPV* = *TP*/(*TP*+*FP*)), Sensitivity (*Se* = *TP*/(*TP*+*FN*)), *F1-score* (*F1* = 2**PPV***Se*/(*PPV*+*Se*)) where: *TP-True Positive*, *FP-False Positive*, *FN-False Negative*.

2.1. Arrhythmia Classifier

The arrhythmia classifier proposed in this paper leverages differences in QRS detection times across multiple QRS detectors as a key feature for distinguishing between various types of heartbeats (Figs. 1 and 2). Different QRS detection algorithms apply different sampling strategies (usually uniform sampling but also the level-crossing sampling [14] as in [10]), distinct signal filtering techniques (linear and non-linear) during the preprocessing stage, and implement unique decision-making rules [15], which naturally result in variations in the R-peak detection times. The heartbeat classification of the proposed classifier is based on comparison of these differences in R-peak detection times.

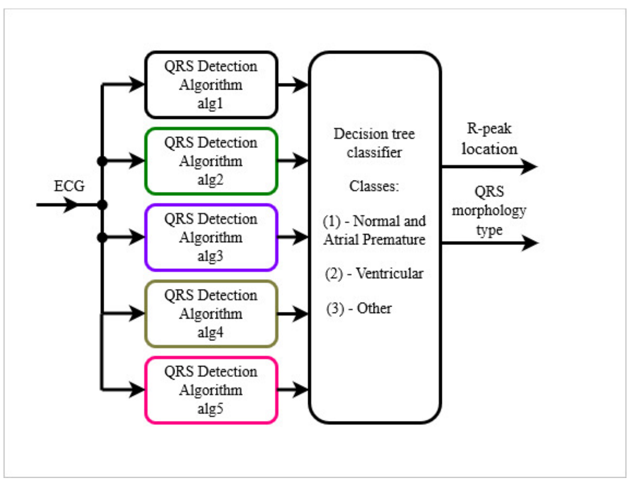

Figure 1. Parallel QRS detectors arrhythmia classifier.

The outputs of the parallel detectors are the inputs to the decision tree classifier. The classification is divided into three QRS morphology types: (1) Normal and Atrial Premature, (2) Ventricular, and (3) Other. The QRS morphology types, Normal and Atrial Premature, were grouped together because their QRS shapes reflect similar electrophysiological ventricular phenomena, resulting in nearly identical ECG traces in lead II, which serves as the basis for classification in this work. The classifier was trained using 70% of the annotations from the MIT-BIH AD dataset, while the remaining 30% were used for testing. Five low-power QRS detectors were selected for the development and testing of the arrhythmia classifier (Fig. 1): alg1 [8], alg2 [9], alg3 [10], alg4 [11], and alg5 [12]. The proposed method was implemented and tested in Python 3.10.4, using the Numpy, Pandas, and Scikit-learn libraries.

2.2. Feature extraction

Feature extraction is crucial to achieving success in arrhythmia classification [3]. Figure 2 presents an excerpt from record 200 of the MIT-BIH AD dataset. It can be observed that for the N-type heartbeat around sample 26400, the QRS detection times vary across different algorithms. Similarly, for the V-type heartbeat around sample 26600, the detection times of the five algorithms also differ.

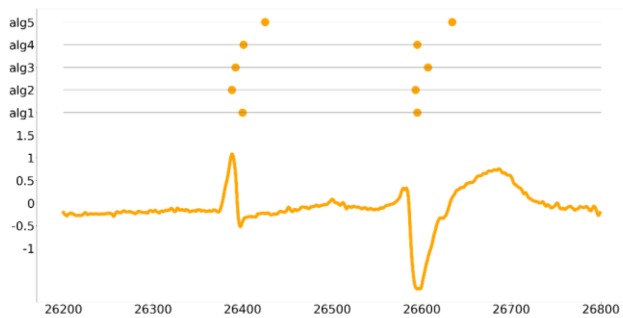

Figure 2. Excerpt from MIT-BIH AD record 200, R-peak detection times of five QRS detection algorithms.

Figure 3. Classification feature for R-peak at time 26400 (Fig. 2) calculated by subtracting R-peak detection time for each of four algorithms from detection time of selected reference algorithm (alg2).

Each heartbeat is associated with five values of the R-peak detection time (expressed as a sample number) from each of the five QRS detection algorithms alg1-alg5. To create the classification feature, one algorithm is chosen as the reference, and the differences in detection time are calculated by subtracting the R-peak detection time of the reference algorithm from the times detected by the other four algorithms (Fig. 3).

In this work, alg2 [9] was chosen as the reference algorithm. The classification feature was analyzed only at time points where alg2 identified an R-peak. If any of the other algorithms failed to detect an R-peak within a ±100-sample range (±278 ms) of alg2 detection time, the data was supplemented by assigning a detection 100 samples earlier than alg2 detection time for that algorithm.

2.2. Decision Tree Classification

The decision tree classifier was selected for this work due to its optimal balance between high classification performance and low computational complexity. Decision tree learning is a supervised approach commonly used in statistics, data mining, and machine learning. The algorithm’s inner workings are straightforward and easy for humans to interpret, and decision trees can be easily visualized. The computational cost of making predictions is logarithmic with respect to the number of data points used to train the tree. In our work, we used Scikit-learn, which implements an optimized version of the CART algorithm [13].

After training, the classification process requires only a limited number of comparisons for each detected R-peak, with the decision tree depth parameter controlling the trade-off between generalization and computational effort. For this work, a tree depth of 12 was selected as the best compromise between accuracy and computational cost. At this depth, the accuracy of the training and test datasets is comparable. Increasing the tree depth leads to overfitting (higher accuracy on the training set but lower on the test set), while reducing the depth lowers both accuracy and computational complexity.

3. Results

Two classifiers were implemented and tested. The first, named Four+One, used alg2 [9] as the reference algorithm, with the differences in detection times of the other four algorithms serving as feature data. The second, a simplified classifier called Two+One, also used alg2 as the reference, but relied on alg3 [10] and alg5 [12] for feature data.

3.1. Four+One arrhythmia classifier

The confusion matrix for the Four+One classifier is shown in Fig. 4, and Table 1 presents the results for Positive Predictivity (PPV), Sensitivity (Se), and F1-score (F1).

Table 1. Results for classifier Four+One for test dataset

Class	*PPV* [%]	*Se* [%]	*F1* [%]	Heartbeats
N+A	93.82	95.51	94.14	23239
V	84.66	76.33	80.28	2112
Other	84.43	79.29	81.79	7639

3.2. Two+One arrhythmia classifier

Similarly, the confusion matrix for the Two+One classifier is displayed in Fig. 5, with the corresponding PPV, Se, and F1 results in Table 2.

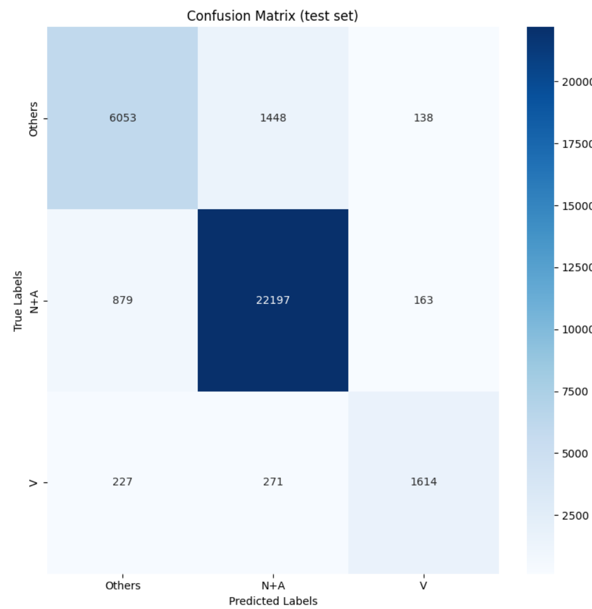

Figure 4. Confusion matrix for Four+One arrhythmia classifier for test dataset.

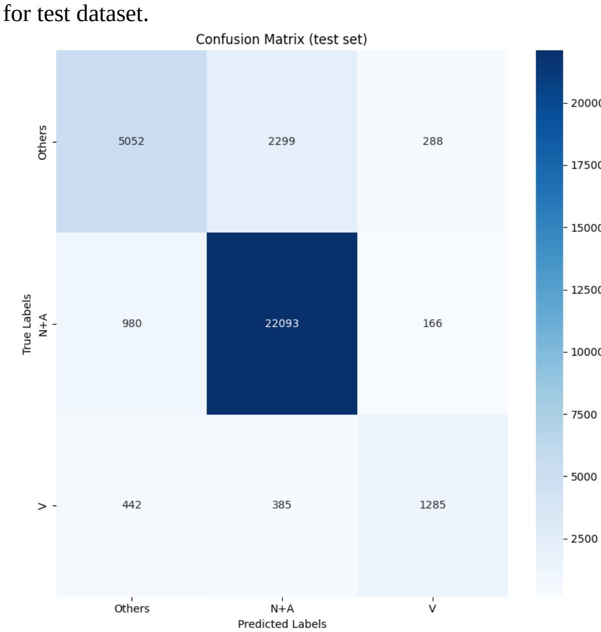

Figure 5. Confusion matrix for Two+One arrhythmia classifier for test dataset.

Table 2. Results for classifier Two+One for test dataset

Class	*PPV* [%]	*Se* [%]	*F1* [%]	Heartbeats
N+A	89.17	95.06	92.02	23239
V	73.89	60.84	66.74	2112
Other	78.02	66.13	71.59	7639

4. Discussion

The overall accuracy of the Four+One classifier is 90.52%, while the Two+One classifier achieves 86.17%. During the training phase, the model maximum depth was set to 12 to prevent overfitting. The accuracy (*ACC*) for the training dataset was 92.03% for the Four+One classifier and 86.74% for the Two+One classifier. The the Four+One classifier utilizes five QRS detectors, while the Two+One classifier uses only three QRS detectors, reducing computational cost at the expense of lower accuracy.

In [3], results from various arrhythmia classifiers report overall accuracy ranging from 83% to 99%. For class V, the PPV ranges from 63% to 99%, while for class N it ranges from 83% to 99%. Sensitivity (Se) for class V spans from 77% to 96%, and for class N, from 80% to 99%.

5. Conclusion

The main contribution of this work is the development of a simple arrhythmia classifier that achieves good classification results comparing to other classifiers reported in the literature, which often require significantly higher computational resources and energy consumption. To examine reducing computational complexity at the expense of lower accuracy, two versions of the classifier were developed and tested, achieving overall classification accuracies of 90.52% for the Four+One classifier, and 86.17% for the Two+One classifier.

Future work will focus on the following areas: exploring additional features to improve discrimination between N and A-type beats, identifying more low-power QRS detection algorithms that, due to their different operating principles, can enhance feature discrimination, further simplifying the computation of the selected QRS algorithms, and training and testing the classifiers on additional databases beyond the MIT-BIH AD dataset.

Acknowledgement

This research was funded by AGH University of Krakow in 2024 as research project no. 16.16.120.773, and by the National Science Centre of Poland under research project DEC-2018/31/B/ST7/03874.

References

[1] L. Neri, I. Corazza, M.T.Oberdier, J. Lago, I. Gallelli, A.F. Cicero, ... C. Borghi, “Comparison between a single-lead ECG garment device and a holter monitor: a signal quality assessment”, *Journal of medical systems*, vol. 48, no. 1, 57, 2024.

[2] N. J. Holter, J. A. Generelli, “Remote recording of physiologic data by radio”, *Rocky Mountain Med J*, pp. 747-751, 1949.

[3] E. J. da S. Luz, W. R. Schwartz, G. Cámara-Chávez, i D. Menotti, ”ECG-based heartbeat classification for arrhythmia detection: a survey”, *Comput Methods Programs Biomed*, vol. 127, pp. 144–164, 2016,

[4] A. Boulif, B. Ananou, M. Ouladsine, S. Delliaux, “A literature review: ECG-based models for arrhythmia diagnosis using artificial intelligence techniques.” *Bioinformatics and Biology Insights*, 17, 2023.

[5] B.-U. Kohler, C. Hennig, R. Orglmeister, “The principles of software QRS detection”, *IEEE Engineering in Medicine and Biology Magazine* 21.1, 42-57, 2002.

[6] M. Elgendi, B. Eskofier, S. Dokos, D. Abbott, “Revisiting QRS detection methodologies for portable, wearable, battery-operated and wireless ECG systems”, *PloS one*, 2014.

[7] A. L. Goldberger, L.A. N. Amaral, L. Glass, J.M. Hausdorff, P. Ch. Ivanov, R.G. Mark ... & H.E. Stanley, “PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals”, Circulation 101 (23), e215-e220, 2000.

[8] R. Gutiérrez-Rivas, J. J. García, W. P. Marnane, Á. Hernández, “Novel real-time low-complexity QRS complex detector based on adaptive thresholding”, *IEEE Sensors Journal*, vol. 15, no. 10, pp. 6036–6043, 2015.

[9] W. Reklewski, K. Heryan, M. Miśkowicz, P. Augustyniak, “Real time ECG R-peak detection by extremum sampling”, *Proceedings - 6th Intern. Conf. on Event-Based Control, Communication and Signal Processing, EBCCSP*, 2020.

[10] N. Ravanshad, H. Rezaee-Dehsorkh, R. Lotfi, Y. Lian, “A level-crossing based QRS-detection algorithm for wearable ECG sensors”, *IEEE Journal of Biomedical Health Informatics*, vol. 18, no. 1, pp. 183–192, 2014.

[11] J. Pan, W. J. Tompkins, “A real-time QRS detection algorithm”, *IEEE Transactions on Biomedical Engineering*, vol. BME-32, no. 3, pp. 230-236, 1985.

[12] Ö. Yakut, E. D. Bolat, “An improved QRS complex detection method having low computational load”, *Biomedical Signal Processing and Control*, vol. 42, pp. 230–241, 2018.

[13] X. Wu, V. Kumar, J. Ross Quinlan, J. Ghosh, Q Yang, H. Motoda, ... D. Steinberg, “Top 10 algorithms in data mining”, *Knowledge and information systems*, vol. 14, pp. 1-37, 2008.

[14] M. Miśkowicz, “Bandwidth requirements for event-driven observations of continuous-time variable”, *IFAC Proceedings Volumes* 37.18, pp. 465-470, 2004.

[15] W. Reklewski, M. Miśkowicz, P. Augustyniak, “QRS detector performance evaluation aware of temporal accuracy and presence of noise”, *Sensors*, vol. 24, no. 5, 1698, 2024.

Address for correspondence:
Piotr Augustyniak
AGH University of Krakow, al. Adama Mickiewicza 30, 30-059 Kraków, Poland, Email: august@agh.edu.pl