Deep Learning End-to-End Approach for Precise QRS Complex Delineation Using Temporal Region-Based Convolutional Neural Networks

Richard Redina[1,2], Jakub Hejc[2,3], Fabian Theurl[4], Tomas Novotny[5], Irena Andrsova[5], Katerina Hnatkova[6], Zdenek Starek[2,7], Marina Filipenska[1], Axel Bauer[4], Marek Malik[5,6]

[1] Department of Biomedical Engineering, Faculty of Electrical Engineering and Communication, Brno University of Technology, Brno, Czech Republic
[2] International Clinical Research Centre, St. Anne's University Hospital, Brno, Czech Republic
[3] Department of Pediatrics, Children's Hospital, University Hospital Brno, Brno, Czech Republic
[4]University Clinic of Internal Medicine III, Medical University of Innsbruck, Innsbruck, Austria
[5] Department of Internal Medicine and Cardiology, Faculty of Medicine, Masaryk University, Brno, Czech Republic
[6] National Heart and Lung Institute, Imperial College London, London, United Kingdom
[7] 1st Department of Internal Medicine, Cardioangiology, St. Anne's University Hospital in Brno and Faculty of Medicine of Masaryk University, Brno, Czech Republic

Abstract

Advancements in clinical diagnosis of heart disease are driven by technological innovations and signal processing developments. ECG segmentation, particularly QRS complex detection, plays a crucial role in cardiac cycle analysis. Deep learning has revolutionized automated ECG analysis, enhancing diagnostic accuracy significantly.

This paper proposes an optimized Region Proposal Network (RPN) architecture for QRS complex detection, specifically designed for 1D signals. Leveraging a vast ECG dataset, extensive data augmentation, feature extraction, and RPN-based QRS detection were employed.

Our method achieved a QRS detection F_1 score of up to 99 %, highlighting its high reliability. Furthermore, QRS complex delineation exhibited deviations typically within 8 ms. The results were verified using a publicly available Lobachevsky University Electrocardiography Database (LUDB), with validation yielding F_1 score of 91.59 % for QRS detection and RMSE of 10.38 ms for QRS complex delineation.

The optimized RPN architecture for QRS complex detection presents a promising solution for efficient ECG analysis.

1. Introduction

The electrocardiogram (ECG) is a fundamental tool for diagnosing various cardiac pathologies. Accurate segmentation of the ECG waveform, particularly the QRS complex, is crucial for rhythm analysis [1] and risk assessment in depolarization disorders. Precise QRS delineation is particularly important in signal-averaged ECG (SA-ECG) methods, where high signal-to-noise ratio (SNR) is achieved through the alignment and aggregation of numerous QRS complexes.

Current QRS delineation methods face several challenges. Conventional expert systems and machine learning (ML) algorithms often rely on hand-crafted rules that may not capture the wide variability of QRS waveforms across different patient cohorts. Furthermore, conventional semantic segmentation approaches require fine-grained temporal features, resulting in high computational overhead and the need for additional post-processing due to non-contiguous region boundaries. Inaccurately detected QRS boundaries necessitate further computationally expensive steps during the alignment process in SA-ECG methods.

This paper proposes a novel end-to-end approach for precise QRS complex delineation based on a temporal region-based convolutional neural network (tR-CNN). R-CNNs [2] are two-stage deep learning (DL) architectures that combine the identification of region candidates followed by boundary refinement and classification.

Prior attempts [3] have explored R-CNNs for QRS detection using an inefficient image-based approach on a 2D signal representation. Herein, we present an efficient R-CNN architecture built upon a ResNet backbone customized for processing the 1D representation of multi-lead ECGs. Our model is designed for fast and accurate QRS delineation, allowing further extension to end-to-end

QRS alignment, morphology-based sorting, and classification. This has the potential to significantly reduce the post-processing required for SA-ECG analysis.

2. Methods

2.1. MediBeat Database

An internal 12-lead ECG database (MediBeat) from 10 917 patients was used for training and model optimization. The basic demographics are summarized in Table 1. The measurements were conducted in accordance with ethical principles. Each patient was characterized by one *representative* – an ECG cycle that was generated as the median of several heartbeats during a 10-second 12 channel ECG measurement. Each *representative* was annotated with QRS onset and offset points.

For the purpose of network training, the dataset was randomly divided into training and test sets in the ratio 50:50. The training set was further divided into 5 folds for the purpose of k-fold cross-validation. Eight uncorrelated channels where used as input.

2.2. Lobachevsky University Electrocardiography Database

Our proposed method was tested on the Lobachevsky University Electrocardiography Database (LUDB) [4]. This publicly available database contains 12-lead ECGs from 200 patients. The main reason why this database was chosen is that all signals contain annotations related to the onset and offset of the QRS complex. The fact that the database contains different types of arrhythmias (I degree AV-block, Left anterior hemiblock) is also an advantage for the verification of correct functionality. The database characteristics are summarized in Table 1.

Table 1. Demographic description of the databases used

Attribute	MediBeat DB	LUDB
Number of subjects	10 917	200
Average age [years]	65	52
Sex: females	5271	85
Length of sample [s]	1.2	10
Sampling frequency [Hz]	500	500

2.3. Feature Extraction

The feature extraction process involves creating a feature map from input data. While many studies in image processing rely on architectures like ResNet-101 [2] for this purpose, our model employs an approach tailored for cardiac electrograms.

Instead of using conventional image processing architectures, we utilized a 1D ResNet specifically developed by the authors for processing cardiac electrograms [5]. This model decomposes the signals in four layers, with increasing the number of filters (8, 56, 104, 152). Additionally, the signals are undersampled by half in each layer, with one layer consisting of six residual blocks, and the output of each layer is processed with an anti-aliasing filter. Subsequently, feature maps undersampled by factor of 16 are utilized for proposal estimation.

The architecture was optimized using Bayesian optimization, ensuring efficient parameter tuning and model performance enhancement.

2.4. Region Proposal Network

A key component of the model is the Region Proposal Network (RPN) [2]. In contrast to semantic segmentation's approach of classifying each ECG sample, a RPN generates a triplet (position $\hat{T}_x$, width $\hat{T}_w$ and objectness score $\hat{Y}$) to define the object:

$$F(\mathbf{X}) \rightarrow (\hat{\mathbf{T}}_{\mathbf{x}}, \hat{\mathbf{T}}_{\mathbf{w}}, \hat{\mathbf{Y}}) \tag{1}$$

First, anchors of different lengths (4, 6, 8 samples) are generated for each position in the feature map. Anchors act as indicators of object presence, with the presence of a QRS complex within an anchor determined by the Intersection over Union (IoU) metric, with a threshold of 0.7. The anchor offsets t_x (2) and t_w (3) are computed relative to a center of reference anchor x_a and its width w_a. The RPN is trained to estimate these offsets.

$$t_x = \frac{x - x_a}{w_a}, \tag{2}$$

$$t_w = \log(\frac{w}{w_a}), \tag{3}$$

The Proposal Module (see Figure 1) consists of two branches — regression and classification. Each branch either infers the required offsets and determines the objectness score for anchors. Loss functions for each branch are defined as follows:

$$L_{cls}(y, \hat{y}) = -\left[y \log(\hat{y}) + (1 - y) \log(1 - \hat{y})\right], \tag{4}$$

$$L_{reg}(t, \hat{t}) = \begin{cases} 0.5(t - \hat{t})^2, & if|t - \hat{t}| < 1 \\ |t - \hat{t}| - 0.5, & otherwise \end{cases} \tag{5}$$

The overall loss $L_B(y, t)$ for the batch B is computed as a weighted sum of classification and regression losses:

$$L_B(\mathbf{y}, \mathbf{t}) = \frac{w_{cls}}{B} \sum_i L_{cls}(y_i, \hat{y}_i) + \frac{w_{reg}}{B} \sum_i L_{reg}(t_i, \hat{t}_i), \tag{6}$$

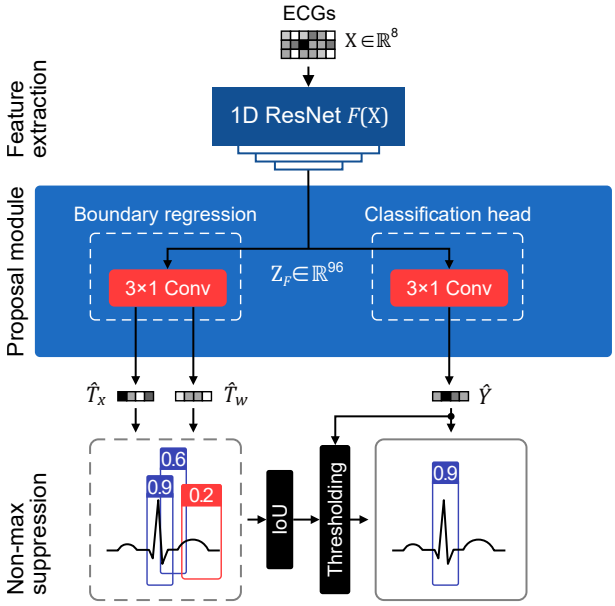

Figure 1. Diagram of the method used. Eight-channel signals (X) are input to the 1D ResNet. The ninety-six channel feature map (Z_F) enters two branches of the proposal box from which bounding boxes indicating the presence of a QRS complex emerge.

where $\hat{y}_i$ represents the objectness score, $\hat{t}_i$ are the offsets for the object obtained from the RPN regression branch, w_{reg} and w_{cls} are weights assigned to the two losses, B denotes the batch size and t_i and y_i are the references. The loss functions L_{cls} (4) and L_{reg} (5) are designed to penalize misclassifications and inaccurate regression predictions, respectively.

Network training was conducted over 140 epochs with an initial learning rate of 0.005, which was decreased every 30 epochs by multiplying it by $\gamma = 0.5$. Training data was organized into batches of 240 samples, and hyperparameters w_{reg} (set to 3.0) and w_{cls} (set to 8.5) were determined using Bayesian optimization.

3. Results and Discussion

A proposal was considered a true positive when the IoU between the predicted boundaries and the reference QS interval exceeded 0.5. Otherwise, the proposal was classified as a false positive. If there was a situation where the QRS complex did not overlap with a single proposal, it was labeled as false negative. From these three metrics, the basic statistical parameters – *recall (Re)*, *precision (Pr)* and F_1 *score* – were calculated.

The position of all true positive proposals were compared with the corresponding reference and *Root Mean Square Error (RMSE)* for the onset and offset of the QRS complex were calculated.

Table 2. QRS instance detection results on MediBeat DB

Dataset	Re [%]	Pr [%]	F_1[%]
Training	99.96±.05	97.88±1.04	98.91± .52
Validation	99.93	99.27	99.60

Table 3. QRS delineation results on MediBeat DB

Dataset	$\text{RMSE}_{on}[ms]$	$\text{RMSE}_{off}[ms]$
Training	8.41±.75	8.69±1.14
Validation	6.72	6.20

Results of training, validation and verification of function at MediBeat DB are captured in Tables 2 and 3, with Table 2 presents the results of instance detection. Table 3 displays the delineation results, evaluating model precision in QRS onset/offset position estimation instances. The training results are given in mean ± standard deviation format as the result of 5-fold cross-validation.

The achieved high $F_1 score$ on both the training and validation datasets is notable. Better performance on the validation set can be probably addressed to the utilization of dataset augmentation techniques, which were varied to maximize signal variability.

The QRS detection results achieving a F_1 of up to 99 % demonstrate the method's high reliability. Additionally, the QRS complex delineation results, with millisecond-level accuracy, indicate precise detection of the QRS complex boundaries.

Table 4. Comparison of QRS detection on LUDB

Author	Re [%]	Pr [%]	F_1[%]
Duraj et al. [6]	92.00	95.00	93.74
Darmawahyuni et al. [7]	95.93	95.94	95.93
Ivora et al. [8]	n/a	n/a	88.70
Proposed study	88.64	94.75	91.59

Table 5. Comparison of QRS delineation on LUDB

Author	$\text{MAE}_{on}[ms]$	$\text{MAE}_{off}[ms]$
Hssayeni et al. [9]	7.30 ± 4.57	9.71 ± 5.64
Proposed study	9.06	10.38

The proposed method underwent independent validation using LUDB, and was compared to existing methods utilizing the same database (see Tables 4 and 5). Considering that LUDB was not used for training our model, a slight decrease in evaluation metrics regarding QRS complex detection is understandable. Similarly to [8], where LUDB was also used for independent validation. The $F_1 score$ decreased by just under 11 percent.

In comparison of QRS complex alignments, our model yields results that are not as favorable as those of the sec-

ond method, as assessed using *mean absolute error (MAE)* following the methodology of Hssayeni et al. Nevertheless, our method's *MAE* falls within $\pm 1\sigma$ of the second method, ensuring the comparability of results.

4. Conclusion

The widely used Region Proposal Network architecture has been successfully implemented for efficient use on 1D signals to detect and delineate QRS complexes.

The main innovation we bring in our presented algorithm is that it is a solution fully focused on 1D signals as opposed to [3, 10].

Our presented method has several advantages. The main advantage of this approach over the semantic signal segmentation used in [5] is the fact that there is no need to decide the membership of each sample of the original signal. Secondly, direct instance detection offers significant advantages by streamlining post-processing compared to semantic segmentation, enabling seamless integration of modules for subsequent tasks.

The presented method proved to be robust both when validated on internal and external database.

Acknowledgments

RR is Brno Ph.D. Talent Scholarship Holder.

This work was supported by a grant for Development of Advanced Methods for Signal, Systems and Data Analysis in Bioengineering and Bioinformatics (FEKT-S-23-8196).

References

[1] Beraza I, Romero I. Comparative study of algorithms for ecg segmentation. Biomedical Signal Processing and Control April 2017;34:166–173. ISSN 1746-8094. URL http://dx.doi.org/10.1016/j.bspc.2017.01.013.

[2] Ren S, He K, Girshick R, Sun J. Faster r-cnn: Towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis and Machine Intelligence June 2017;39(6):1137–1149. ISSN 2160-9292. URL http://dx.doi.org/10.1109/TPAMI.2016.2577031.

[3] Yang H, Huang M, Cai Z, Yao Y, Liu C. A faster r cnn-based real-time qrs detector. In Computing in Cardiology Conference (CinC), CinC2019. Computing in Cardiology. ISSN 2325-887X, December 2019; URL http://dx.doi.org/10.22489/CinC.2019.053.

[4] Kalyakulina A, Yusipov I, Moskalenko V, Nikolskiy A, Kosonogov K, Zolotykh N, Ivanchenko M. Lobachevsky university electrocardiography database, 2021. URL https://physionet.org/content/ludb/1.0.1/.

[5] Hejc J, Pospisil D, Novotna P, Pesl M, Janousek O, Ronzhina M, Starek Z. Segmentation of atrial electrical activity in intracardiac electrograms (iecgs) using convolutional neural network (cnn) trained on small imbalanced dataset. In 2021 Computing in Cardiology (CinC). IEEE, September 2021; URL http://dx.doi.org/10.23919/CinC53138.2021.9662729.

[6] Duraj K, Piaseczna N, Kostka P, Tkacz E. Semantic segmentation of 12-lead ecg using 1d residual u-net with squeeze-excitation blocks. Applied Sciences March 2022;12(7):3332. ISSN 2076-3417. URL http://dx.doi.org/10.3390/app12073332.

[7] Darmawahyuni A, Nurmaini S, Rachmatullah MN, Avi PP, Teguh SBP, Sapitri AI, Tutuko B, Firdaus F. Improved delineation model of a standard 12-lead electrocardiogram based on a deep learning algorithm. BMC Medical Informatics and Decision Making July 2023;23(1). ISSN 1472-6947. URL http://dx.doi.org/10.1186/s12911-023-02233-0.

[8] Ivora A, Viscor I, Nejedly P, Smisek R, Koscova Z, Bulkova V, Halamek J, Jurak P, Plesinger F. Qrs detection and classification in holter ecg data in one inference step. Scientific Reports July 2022;12(1). ISSN 2045-2322. URL http://dx.doi.org/10.1038/s41598-022-16517-4.

[9] Hssayeni MD, Andalib A, Singh R, Pava D, Li K, Borzak S, Chait R, Kale K. Ecg fiducial point localization using a deep learning model. In 2022 21st IEEE International Conference on Machine Learning and Applications (ICMLA). IEEE, December 2022; URL http://dx.doi.org/10.1109/ICMLA55696.2022.00052.

[10] Yu R, Gao Y, Duan X, Zhu T, Wang Z, Jiao B. Qrs detection and measurement method of ecg paper based on convolutional neural networks. In 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, July 2018; URL http://dx.doi.org/10.1109/EMBC.2018.8513132.

Address for correspondence:

Richard Redina,
Department of Biomedical Engineering,
Brno University of Technology,
Technicka 3058/12, Brno, 61200,
Czech Republic
E-mail adress: 195715@vut.cz