

Automated ECG Segmentation in Telehealth System: A Case Study with Real Brazilian Data

Matheus Felipe Akira de Assis Oliveira¹, Jermana Lopes Moraes², Derick Matheus Oliveira¹,
Wagner Meira Jr¹, Antonio Luiz Pinho Ribeiro¹

¹ Universidade Federal de Minas Gerais, Belo Horizonte, Brazil

² Universidade Federal do Ceará, Sobral, Brazil

Abstract

The segmentation of electrocardiogram (ECG) signals is essential for accurately analyzing critical features such as the P wave, the QRS complex, and the T wave, aiding in the diagnosis of heart disorders. This study explores Deep learning techniques applied to data from Telehealth Network of Minas Gerais (TNMG). The model achieved a precision of 0.88 (P wave), 0.91 (QRS), and 0.93 (T wave), with recall of 0.81 (P wave), 0.91 (QRS), and 0.85 (T wave), and F1-scores of 0.84 (P wave), 0.91 (QRS), and 0.84 (T wave), resulting in an overall accuracy of 91.14%. With the Hospital das Clínicas in Belo Horizonte processing thousands of ECG exams daily, automated segmentation can speed up triage and improve decision-making, contributing to faster and more efficient diagnoses within the Brazilian public health system.

1. Introduction

The electrocardiogram (ECG) is the most basic and important tool to diagnose heart diseases, which remain one of the leading causes of death in 2024 [1]. In many countries, cardiology services are predominantly concentrated in urban centers. Consequently, residents of suburban and rural regions often face limited or no access to essential cardiac evaluations, such as ECG [2]. In such scenario, the Telehealth Network of Minas Gerais (TNMG) played a fundamental role, which serves with telecardiology services more than 1200 cities throughout Brazil [3].

Currently, TNMG receives approximately 5000 12-lead ECG recordings to diagnose every day [4]. One of the most important steps in cardiac diagnosis is the measurement of the waves and intervals from the ECG recordings. Therefore, an automated segmentation of important ECG components, such as the P waves, the QRS complex, and the T wave, is crucial to improving diagnostic accuracy, especially in Telehealth systems.

Several state-of-the-art algorithms for ECG segmenta-

tion have been introduced in the literature. Most of these algorithms are based on classical machine learning, deep learning models, and digital signal processing techniques. Among the most well-known models is the U-Net, widely applied in biomedical segmentation, utilizing an encoder-decoder architecture that enables precise segmentation of ECG waves [5]. The Convolutional LSTM (ConvLSTM) combines convolutional layers and LSTM, capturing spatial and temporal features of ECG signals, outperforming traditional methods [6]. Another relevant model is the Attention-based Conv-BiLSTM, which employs a bidirectional LSTM architecture with attention mechanisms, enabling precise segmentation and fiducial point detection [7].

We believe that these models currently available in the literature present a key limitation regarding to the applicability to the TNMG system: all of them were developed and validated based on datasets not representative of a real-world clinical application, in which the types of cardiac alterations are insufficiently represented.

To fill this gap, we used the SegNet architecture [8] in conjunction with post-processing technique to automatically segment the P, QRS, and T waves of the 12-lead ECG of a real-world medical application.

We selected the SegNet model due to its effectiveness in segmenting temporal signals and handling variations in biomedical data. Its architecture allows for the precise detection of key features in ECG signals, in addition to being memory-efficient and robust against noise, making it suitable for clinical applications and ECG segmentation.

The Segnet model was tested by means of different performance metrics using a dataset of 12-lead ECG recordings from TNMG (504 ECG recordings with the 30 main types of cardiac abnormalities present in the TMNG).

By adapting the SegNet model [8] to these specific data, our goal was to demonstrate the feasibility and effectiveness of automated segmentation techniques in a real-world scenario, where of the ECG recordings are representative of the clinical reality in terms of types and occurrence of

pathological recordings.

2. Materials and methods

2.1. Dataset

For this study, we used a 12-lead ECG dataset, which was provided by TNMG (approval by Research Ethics Committee, protocol: 68496317.7.0000.5149). Each dataset consists of 12-lead ECG recordings, length of 10 s, sampling frequency $\leq 300\text{Hz}$, resolution $\leq 12\text{bits}$.

The dataset used consists of 504 ECG recordings, annotated by a specialist from TNMG and distributed into 30 classes, with 29 cardiac abnormalities and healthy ECG recordings. The annotations were made in four specific leads: I, aVF, V2, and II. These leads were chosen due to the orthogonality of the first three, and the importance of lead II for clinical analysis for the TNMG specialists.

The data was randomly divided, with 70% allocated for training, 15% for validation and 15% for testing, ensuring an appropriate balance between classes. Among the cardiac anomalies present are left atrial overload, ventricular extrasystoles, first-degree atrioventricular block (BAV1), atrial fibrillation (AF), among others.

The methodology applied in this study is composed of three main stages: preprocessing, segmentation, and extraction of the segmented signal. Each of these stages is detailed in the following subsections.

2.2. Preprocessing

Each ECG lead underwent a series of transformations aimed at improving signal quality for subsequent analysis. Initially, signals with varying sampling frequencies were resampled to 400 Hz using the code available in [9]. Then, a baseline wander removal technique [10] was applied, combined with a high-pass filter with a cutoff frequency of 0.8 Hz to remove low-frequency components that could interfere with signal analysis. The filter was configured to attenuate frequencies below 0.2 Hz with an attenuation of 40 dB, while within the allowed range (above 0.8 Hz), the signal variation was minimal (ripple of 0.5 dB).

Then, the signal was normalized, ensuring that the P, T waves, and the QRS complex were clearly preserved for the segmentation stage.

2.3. SegNet

SegNet is a convolutional neural network initially designed for image segmentation [8], but its flexibility allows it to be applied to other fields, such as ECG signal segmentation. Its efficient architecture enables the segmentation of key components in ECG signals, such as the P and T waves and the QRS complex, while preserving essential

information to ensure accurate segmentation. By adapting SegNet for ECGs, this architecture can automatically identify critical intervals within the cardiac signals, assisting in clinical analysis and diagnosis.

The architecture of SegNet for ECGs retains its two main parts: the encoder and the decoder (Figure 1). The encoder, using convolutional layers, extracts features from the ECG signals, compressing this data to highlight essential information like rhythmic and electrical conduction patterns in the heart. The decoder then reconstructs this information into a segmented format, making it easier to visualize and interpret the intervals of the P, T waves, and QRS complexes. In this process, pooling indices from the encoding phase are reused during unpooling, ensuring spatially precise segmentation of the signals.

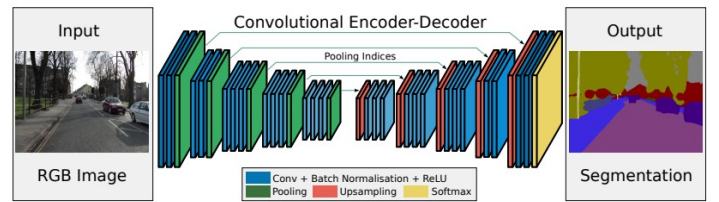


Figure 1. Architecture of SegNet for ECG signal segmentation.

For the training, the individual leads (I, II, aVF, and V2). Also, the training was conducted over 500 epochs, using a batch size of 32 and the Adam optimizer. The SegNet architecture was configured with 32 layers and an output layer containing four neurons, corresponding to the four main ECG recordings categories: P waves, T waves, QRS complexes, and the baseline.

2.4. Extraction of the Segmented Signal

The final stage involves extracting the segmented signal. Here, each ECG lead was processed individually to obtain the markings of the identified waves, including the baseline, which were exported for further analysis. The extracted markings include the temporal boundaries of each segmented wave (P, T, QRS) and the baseline, allowing detailed clinical analysis of each lead. This segmented information is used to generate insights into potential cardiac conditions of the patients, ensuring that the analysis reflects the specific characteristics of each ECG derivation.

3. Results

The model results were compared with the annotations made by the specialist from TNMG, using the leads I, II, aVF, and V2. We compared our model with the ground truth, i.e., manual labeling by an expert. Positive (P) refers

to correctly identified ECG segments, and negative (N) refers to segments incorrectly classified or not identified. First, we computed the true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). Next, we calculated the following metrics: accuracy, precision, F1-score, and recall.

Then, we evaluated the overall accuracy of the model, which summarizes the performance across the entire dataset. In addition, we analyzed precision, F1-score, and recall, calculated separately for the P waves, QRS complex, T waves, and baseline segments. Figure 2 shows the confusion matrix obtained from the test data for each class (P waves, QRS complex, T waves, and baseline).

Figure 2 shows the confusion matrix obtained from the test data.

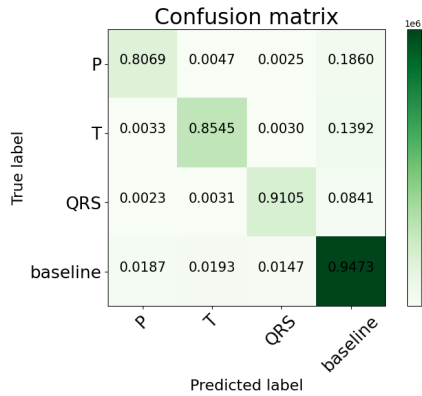


Figure 2. Confusion Matrix for the P, QRS, T waves, and Baseline.

After examining the confusion matrix, we computed the performance metrics. Table 1 provides the overall accuracy of the model, along with the precision, F1-score, and recall values for the P waves, QRS complex, T waves, and baseline.

Metric	P Wave	QRS Complex	T Wave	Baseline
Precision	0.88	0.91	0.93	0.91
Recall	0.81	0.91	0.85	0.95
F1-score	0.84	0.91	0.84	0.93
Overall Accuracy	0.9114			

Table 1. Performance metrics for the P waves, QRS complex, T wave, and baseline.

Additionally, Bland-Altman plots were generated to assess the agreement between the model's predictions and the expert annotations for the P waves, QRS complex, and T waves. In these plots, the 'x' axis represents the expert annotations, and the 'y' axis represents the model's predictions, both in milliseconds. Figures 3, 4, and 5 show the Bland-Altman plots for the P wave, QRS complex, and T wave, respectively.

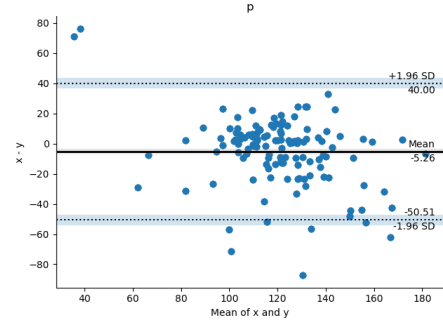


Figure 3. Bland-Altman plot for the P wave.

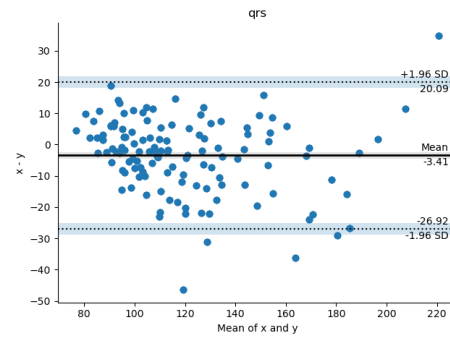


Figure 4. Bland-Altman plot for the QRS complex.

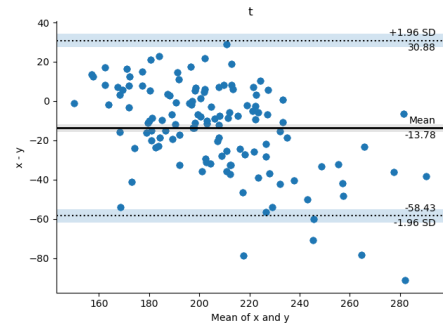


Figure 5. Bland-Altman plot for the T wave.

4. Discussion

Although the SegNet model used in this work has a relatively simple architecture, it is widely established in the literature for segmentation tasks, especially in medical contexts [11]. The simplicity of the model does not compromise its effectiveness, as it has been fine-tuned to process Brazilian ECG signals, achieving consistent and robust results [12]. This suggests that there is not necessarily a need to use more complex or sophisticated models when a simpler, well-established model can meet the application's requirements. In many cases, the effectiveness of a segmentation model lies less in its complexity and more in

its adaptation to the local clinical reality, which reinforces the value of an evidence-based approach and fine-tuning of known models.

A major challenge that stands out is the need for datasets that adequately reflect the local clinical reality. The use of data that do not capture the specific characteristics of the population being served, such as local cardiac anomalies may limit the generalization and applicability of the model in a real-world clinical application [13]. Therefore, greater data representativeness would be ideal to increase the model's accuracy, especially concerning cardiac anomalies that occur in the local population. Creating datasets with a representative number of cardiac anomalies, associated with specific genetic and physiological traits, can significantly improve the performance of segmentation models and their clinical applicability.

Thus, it is crucial to direct efforts not only towards developing new models but also towards creating robust and representative datasets. Building databases that capture the diversity and particularities of the local clinical reality, such as cardiac patterns and anomalies, would be a crucial step towards improving the performance of relatively simple models, making them effective for clinical applications in Brazil and other regional contexts.

4.1. Conclusion

Despite the limited data available for training the model, the results demonstrate that it was effective in segmenting the P waves, QRS complex, T wave, and baseline with good precision. The statistical metrics, along with the visual analysis in the Bland-Altman plots, indicated that the model handled the challenges posed by noisy data and the variation in the Brazilian population well. The model's overall accuracy was satisfactory, and the precision, F1-score, and recall metrics reinforce the robustness of the model, especially considering the limited data conditions available for training.

Acknowledgments

This work was partially funded by CNPq and CAPES, as well as CIIA-Saúde and IAIA - INCT on AI. We also thank the Telehealth Center of Minas Gerais for access to the data and the productive discussions within the scope of this work.

References

- [1] WH F. World heart report 2024, 2024. URL <https://worldheart-federation.org/resource/world-heart-report-2024>. Accessed: 2024-09-24.
- [2] de Moraes JL, Paixão GM, Gomes PR, Mendes EM, Ribeiro ALP, Beda A. A novel algorithm to assess the quality of 12-lead ecg recordings: validation in a real telecardiology application. *Physiological Measurement* 2023; 44(3):035006.
- [3] Ribeiro ALP, Paixao GM, Gomes PR, Ribeiro MH, Ribeiro AH, Canazart JA, Oliveira DM, Ferreira MP, Lima EM, de Moraes JL, et al. Tele-electrocardiography and bigdata: the code (clinical outcomes in digital electrocardiography) study. *Journal of electrocardiology* 2019;57:S75–S78.
- [4] de Telessaúde Hospital das Clínicas UFMG C. Resultados julho 2024. <https://telessaude.hc.ufmg.br/resultados/resultados-julho-2024/>, 2024. Acessado em: 26 Set. 2024.
- [5] Moskalenko V, Zolotykh N, Osipov G. Deep learning for ecg segmentation. *arXiv preprint arXiv200104689* 2020;.
- [6] Malali A, Hiriyannaiah S, G.M. S, K.G. S, N.T. S. Supervised ecg wave segmentation using convolutional lstm. *ICT Express* 2020;6(3):166–169.
- [7] Haleem MS, Pecchia L. A deep learning based ecg segmentation tool for detection of ecg beat parameters. In *2022 IEEE Symposium on Computers and Communications (ISCC)*. IEEE, 2022; 1–5.
- [8] Badrinarayanan V, Kendall A, Cipolla R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *arXiv preprint arXiv151100561* 2015; Last revised Oct. 2016.
- [9] Ribeiro AH. ecg-preprocessing. <https://github.com/antonior92/ecg-preprocessing>, 2022. Acessado em: 02 Fev. 2022.
- [10] Gupta P, Sharma KK, Joshi SD. Baseline wander removal of electrocardiogram signals using multivariate empirical mode decomposition. *Healthcare Technology Letters* Dec. 2015;2(6):164–166. Received: July 19, 2015; Revised: October 2, 2015; Accepted: October 26, 2015.
- [11] Alqazzaz S, Sun X, Yang X, Nokes L. Automated brain tumor segmentation on multi-modal mr image using segnet. *Computational Visual Media* 2019;5(2):209–219.
- [12] Hua J, Chu B, Zou J, Jia J. Ecg signal classification in wearable devices based on compressed domain. *PLoS One* 2023;18(4):e0284008.
- [13] de Moraes JL, Paixão GM, Gomes PR, Mendes EM, Ribeiro ALP, Beda A. A novel algorithm to assess the quality of 12-lead ecg recordings: validation in a real telecardiology application. *Physiological Measurement* 2023; 44(3):035006.

Address for correspondence:

Matheus Felipe Akira de Assis Oliveira
DCC–ICEx–UFMG - Pampulha Belo Horizonte, MG 31270-901
Brazil
matheusfelipeakira@gmail.com