Enhancing ECG Analysis with a Hybrid Deep Learning Approach: Automatic Detection of Significant Features

Enhancing ECG Analysis with a Hybrid Deep Learning Approach: Automatic Detection of Significant Features

Federico Delrio, Valentina De Vitis, Silvia Caligari, Vincenzo Randazzo, Giansalvo Cirrincione, Eros Pasero

Abstract Cardiovascular diseases (CVDs) are the main cause of death worldwide, with arrhythmias posing significant risks. The interpretation of ECG signals requires professional expertise and time, which has limited its accessibility to the general public. This research introduces a novel hybrid deep learning technique for the automatic detection of key features (QRS complex, P and T waves) in electrocardiogram (ECG) signals. The method integrates two Convolutional Neural Networks (CNNs) and two Bidirectional Long Short-Term Memory Networks (BiLSTMs) to form a 2D-CNN-BiLSTM model. The model is trained and evaluated using the "QT Database", a publicly available dataset. ECG signals are first downsampled and then pre-processed using a local linear regression technique to remove the baseline, and the Continuous Wavelet Transform (CWT) to convert 1D ECG signals into 2D images. The model demonstrates superior classification performance in detecting P waves, QRS complexes, and T waves, compared to existing methods. The suggested approach shows promise for improving the precision and effectiveness of ECG analysis and advancing automated cardiac diagnostic techniques.

Key words: ECG; Deep Learning; Arrhythmias; Cardiovascular diseases; Neural Networks; CNN; BiLSTM; Scalogram; Continuous Wavelet Transform; CWT

Federico Delrio (✉) · Valentina de Vitis · Silvia Caligari · Vincenzo Randazzo (✉)· Eros Pasero
DET - Department of Electronics and Telecommunications, Politecnico Di Torino, Turin, Italy, e-mail: federico.delrio@polito.it, vincenzo.randazzo@polito.it

Giansalvo Cirrincione
Lab. LTI, Université de Picardie Jules Verne, Amiens, France; University of South Pacifc, Suva, Fiji e-mail: nimzoexin59@gmail.com

1 Introduction

Cardiovascular diseases (CVDs) are one of the leading causes of mortality worldwide, making prevention and early treatment crucial. One of the most effective tools for diagnosing is the electrocardiogram (ECG), a non-invasive test that records the electrical activity of the heart over time, providing important information in terms of waveform and time intervals. The most important features in ECG signals are the P wave, the QRS complex, and the T wave, which represent the most significant phases of the cardiac cycle [1]. However, the interpretation of ECG signals requires professional expertise, which has limited their accessibility to the general public.

Recent advances in machine learning [2] have shown promising results in medical applications [3], such as gene expression [4] and cancer [5] analysis, risk stratification of Short-QT and Brugada syndromes [6, 7], and multi-omics integration [8, 9, 10]. Deep learning has shown promise in automating the interpretation of ECG signals, improving accuracy and time management in feature detection [11]. This article aims to further explore the potential of deep learning in automating the interpretation of ECG signals, focusing on the development of a method for the automatic detection of relevant features in ECG signals. Indeed, deep learning has increasingly been used in ECG interpretation research, notably for automatic diagnosis of cardiac arrhythmias and blood pressure estimation from ECG signal [12, 13, 14]. Additionally, digitization of ECG records via deep learning, as proposed by [1], enables easy access and storage of ECG reports. A 1D convolutional neural network has been developed for automatic arrhythmia classification with high accuracy [15]. Lastly, a comparative study indicated that nonlinear features carry crucial QRS-complex information, resulting in high classification rates with minimal features [16].

2 Methodology

The methodology employed in this research is a comprehensive, multi-step process that involves data pre-processing, wavelet transformation, and data segmentation. Each of these steps is crucial in the development of a deep learning model for the identification of significant characteristics in electrocardiogram (ECG) signals.

The study uses the QT Database [17, 18] for its diverse, relatively clean ECG signals from various patient categories, including ambulatory recordings, exercise tests, and long-term recordings ranging from 14 to 22 hours. Expert-annotated waveform boundaries help ensure data accuracy. The pre-processing stage uses the Waveform Database tool (WFDB) [19], from PhysioNet [18], to extract recordings, carefully excluding irrelevant data and removing non-annotated zones to maintain study reliability.

2.1 Baseline Elimination

Baseline elimination is another crucial step in the analysis of ECG signals. Baseline wander, a low-frequency noise, can interfere with the interpretation of the ECG signals. Therefore, it is essential to remove this noise to obtain clean and accurate signals. The elimination of baseline wander is accomplished through the implementation of a local linear regression technique. In this approach, a linear regression is computed for a window of approximately 1.5 seconds. The midpoint of this window serves as the reference baseline. Subsequently, the window is shifted by one point, and the process is reiterated. The algorithm in use, as illustrated in Fig.1, is successful in reducing baseline instability. By allocating lesser significance to the anomalies within the regression data, the strategy guarantees that the essential components of the ECG signal, namely the P and T waves and QRS complexes, are emphasized as exceptions in the regression analysis. Consequently, this preserves these vital characteristics while adeptly removing baseline inconsistencies.

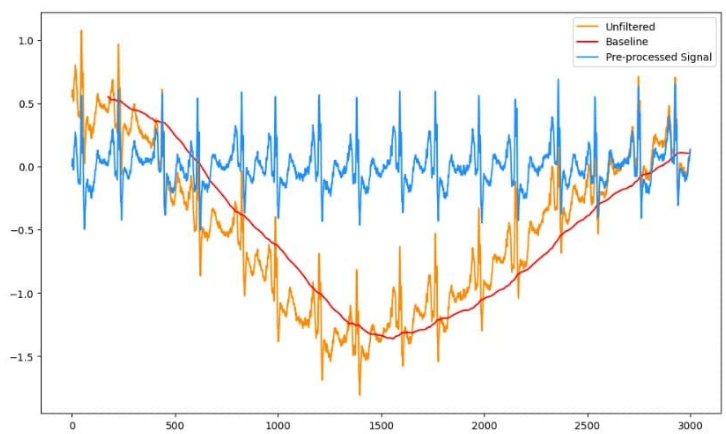

Fig. 1: Example of baseline removal

2.2 Continuous Wavelet Transform and Scalogram Derivation

The Wavelet Transform [20] serves as a mathematical technique employed to break down a signal into its constituent elements. Within this research, the Continuous Wavelet Transform (CWT) [21] is utilized to generate the scalogram, which visually depicts the wavelet transform of the signal.This technique transforms one-dimensional ECG signals into two-dimensional visual representations, thus providing an enhanced perspective of the signal's characteristics. As part of this process, the signal undergoes downsampling from 250 Hz to 125 Hz. This downsampling step reduces the computational complexity of the model, enhancing its efficiency. The CWT takes the form of a matrix comprising wavelet coefficients that are localized according to their scale and position. These coefficients contain valuable insights into the signal's characteristics across various scales and positions.

The scalogram functions as an instrument for constructing and displaying the two-dimensional spectrum obtained via the Continuous Wavelet Transform, as shown in Fig.2 (bottom). This displays the CWT coefficients' absolute magnitudes, providing a visual map of the signal's frequency content. In the scalogram, the range of colors translates to the varying amplitudes of the signal's frequency elements, with darker hues signifying higher amplitudes and paler hues denoting lower ones.

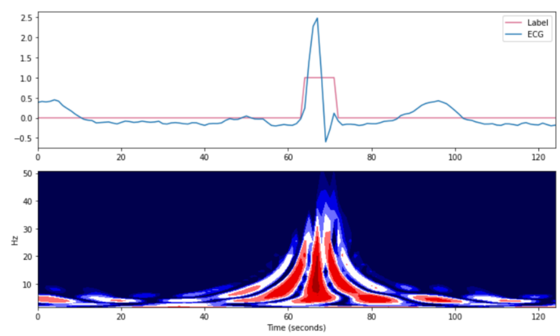

Fig. 2: Top: segmentation example of ECG with relative label for a QRS complex (red). Bottom: 2D Scalogram obtained through CWT method from the ECG.

2.3 Data Segmentation

Data segmentation is a critical process when dealing with lengthy signals that could potentially degrade the performance of dynamic deep learning models such as Convolutional Neural Networks (CNNs). To prevent this degradation, the ECG signals and their related labels are segmented. This process ensures that the CNNs used for feature extraction are fed with manageable data sizes, enhancing the overall performance of the model. This study employs a method of dividing each ECG sample into segments consisting of 125 data points. To maintain continuity, there is an overlap of 10 data points between adjacent segments. This segmentation technique aims to create manageable portions of data suitable for processing by a deep neural network. The two leads in the ECG recording are separated into distinct segments, and both are utilized as features. These segments, comprising 125 samples, serve as input for the deep neural network.

The purpose of the segmentation process extends beyond the mere division of ECG signals into smaller segments; it also ensures that each segment encompasses significant information. Specifically, each window encompasses either a complete QRS complex or a partial QRS complex along with its associated label. This approach guarantees that each segment contains an entire waveform, which is crucial for accurately analyzing the ECG signal. The total number of obtained segments or windows amounts to 15089. An example of segmentation with relative scalogram can be seen in Fig.2(top).

3 Deep learning Architecture

The deep learning architecture employed in this research is a composite of two CNNs and two Bidirectional Long-Short Term Memory Networks (BiLSTMs). This architecture was constructed utilizing Python and TensorFlow and underwent training on a virtual server furnished by Google Colab, which provided a high-performance GPU for efficient training. The primary roles of the CNNs in this project encompass the extraction of signal characteristics from the data. To move from CNN to LSTM, a necessary step is the implementation of time-distributed flattening. This process transforms a two-dimensional matrix into a one-dimensional vector, which can then be fed into a densely connected neural network layer. Long Short-Term Memory (LSTM) networks are capable of forecasting imminent QRS complexes or P-T waves by learning from preceding sequences, thanks to their architecture, which is specifically tailored to manage the temporal characteristics found in ECG data.

The model comprises of dual LSTM layers, each accompanied by a dropout layer with a drop rate of 25%. The model produces a binary output, indicating the presence (1) or absence (0) of a specific feature. Then, the output passes through another layer with time-distributed wrapping, generating one predicted value for each time-step. Following this, two additional layers are incorporated for Batch Normalization and Dropout Regularization (with a 25% rate). Finally, a dense layer employing the softmax activation function is utilized to predict the ultimate binary value. The categorical cross-entropy loss is used in this model, and the Adaptive Moment Estimation, also known as Adam, optimizer was selected because it provides good computational efficiency. The overall model structure can be seen in detail in Fig.3.

4 Results

The QRS complex, P wave, and T wave detection follow the same outlined procedures, methodologies, and model structure mentioned earlier. To enhance the training process, a 10-fold Cross Validation (CV) technique is used. For improving accuracy, loss, and AUC assessment, each of the 10 folds in the CV contains a distinct validation dataset composed of randomly selected samples from the overall dataset.

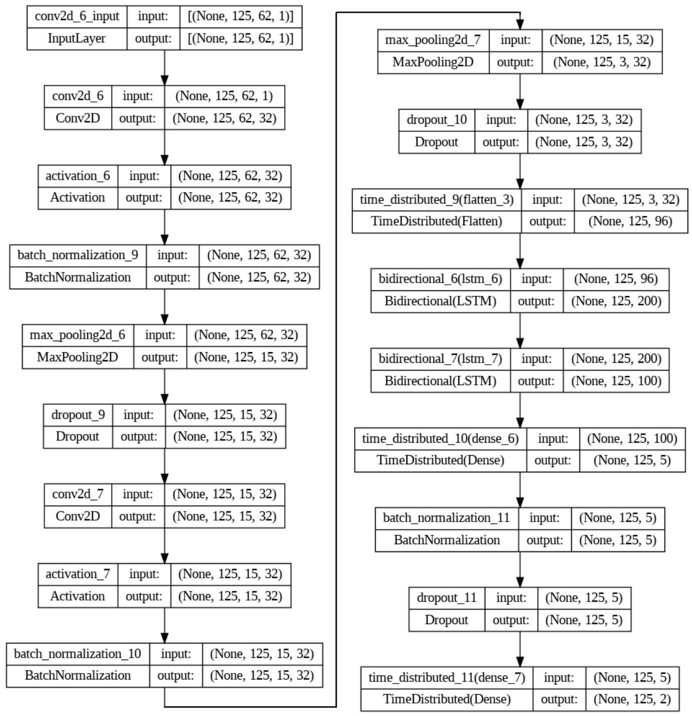

Fig. 3: Deep Neural Network Architecture

Evaluation metrics such as ROC, sensitivity, specificity, and F-score are calculated accordingly. During the training of the model, the categorical cross entropy serves as the chosen loss metric throughout the 30 epochs. For all three networks (QRS,P and T) the average of the 10 k-fold has been calculated, and the final results can be seen in Table1. In particular, the model demonstrates remarkable accuracy in identifying specific elements of ECG signals, such as P, QRS, and T waves, achieving validation accuracy of 0.962, 0.977, and 0.964, respectively.

Table 1: Average scores for all folds for all the analyzed features

	Accuracy	Loss	Sensitivity	Specificity	F1-score
P wave	0.962	0.1228	0.878	0.974	0.8557
QRS complexe	0.977	0.0726	0.923	0.981	0.9126
T wave	0.964	0.1174	0.934	0.974	0.9247

Finally, Fig. 4 reports the losses against the training epochs of the neural networks for QRS complex, P waves, and T waves. The figures employ varying X and Y scales to accommodate the diverse features learned by the three networks, which consequently affects the magnitude of the loss observed in different scenarios. For the purposes of clarity and to more effectively illustrate the trend of the loss function, the figures are limited to approximately 16 epochs, rather than the full 30. This decision was made to highlight the decreasing trend of the loss function, noting that the loss function exhibits minimal changes beyond the 15th epoch. Example of predictions on random test data can be seen in Fig. 5 for QRS, Fig. 6 for P waves and Fig. 7 for T waves.

Furthermore, to showcase the superior performance of the proposed model, a comparative analysis with existing and widely utilized models in this field was conducted. Table 2 presents these results in comparison with other methodologies, including ECG-SegNet [22] and Hidden Markov Models (HMM) [23]. Furthermore, the proposed model shows superior sensitivity and specificity in the detection of these signal components.

Table 2: Comparison of results with different methods

Methods	**P wave**	**QRS Complex**	**T wave**
2D-CNN-BiLSTM	0.962	0.971	0.965
ECG-SegNet (Bidirectional LSTM)	0.920	0.940	0.920
HMM on raw ECG data	0.551	0.790	0.560
HMM on wavelet encoded ECG	0.742	0.944	0.882

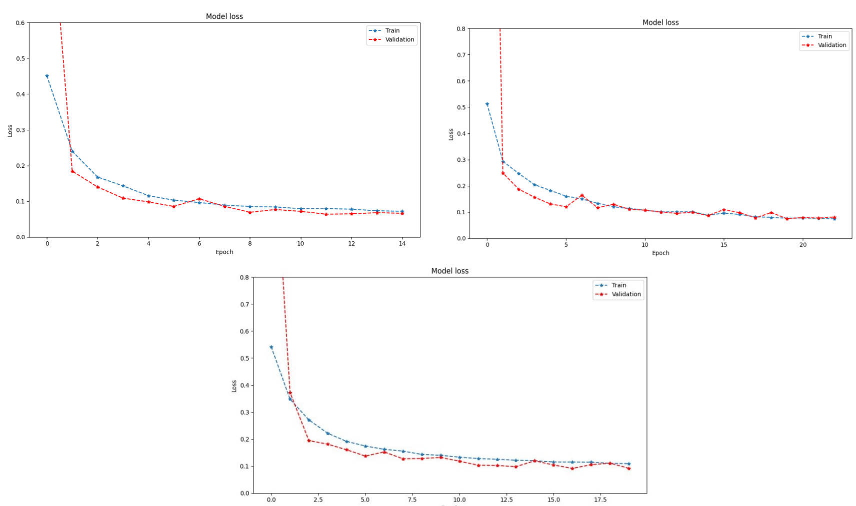

Fig. 4: Training loss (blue) and validation loss (orange) for QRS (top left), P wave (top right), and T wave (bottom) networks.

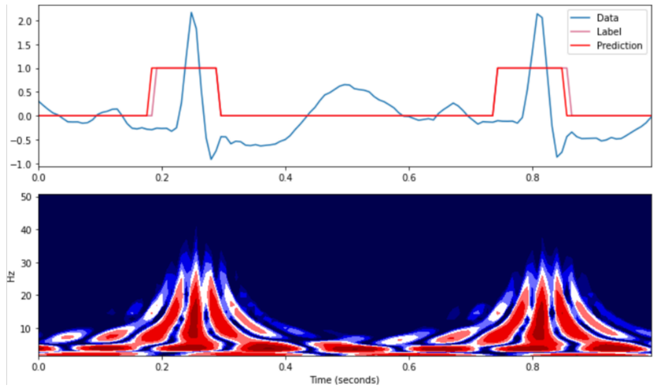

Fig. 5: Example of prediction of QRS complexes with the relative original label

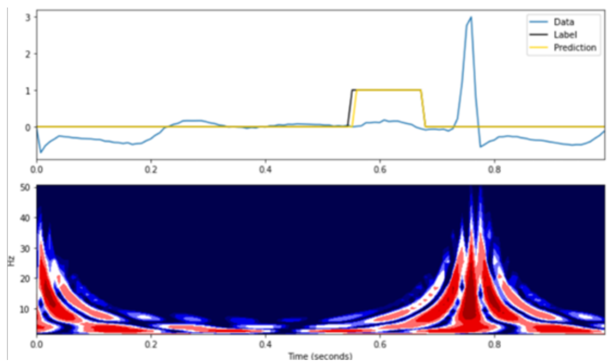

Fig. 6: Example of prediction of P wave with the relative original label

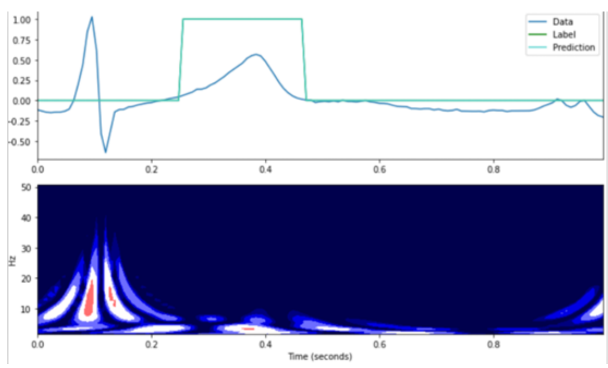

Fig. 7: Example of prediction of T wave with the relative original label

5 Discussion and Conclusion

The aim of this work is to develop an automated system for the analysis of electrocardiogram (ECG) data. The existing methods for ECG analysis rely heavily on human experts and consume a considerable amount of time. To address this issue, the proposed solution utilizes a novel 2D-CNN-BiLSTM model that extracts features from ECG signals. The model combines convolutional neural networks (CNNs) and bidirectional long short-term memory (LSTM) networks. CNNs are employed to extract relevant features, while LSTM networks effectively capture temporal relationships within the ECG data. By applying a pre-processing technique known as CWT, the ECG signals are transformed into 2D scalogram images, which are subsequently segmented into windows and used to train the CNN-BiLSTM model.

The use of this automated system has a lot of potential for optimizing massive screening procedures and automating the diagnosis of cardiac diseases. It can lessen the workload for cardiologists and make it easier to precisely identify ECG features while performing routine monitoring. This automated algorithm can significantly improve patient care and boost the effectiveness of cardiovascular disease diagnosis and prevention by utilizing cutting-edge deep learning techniques. Overall, the proposed hybrid deep learning technique showed promise in automating ECG signal interpretation and improving the precision and effectiveness of ECG analysis. The early detection of cardiovascular diseases may be improved and automated cardiac diagnostic methods may advance thanks to the model's capacity to identify important features in ECG signals. Further research and validation on larger and more diverse datasets are needed to fully assess the performance and generalizability of the proposed method.

Acknowledgments. This research is part of the BrS-AI-ECG project, funded by the Italian Ministry of Foreign Affairs and International Cooperation. Dr. Randazzo acknowledges funding from the research contract no. 32-G-13427-2 (DM 1062/2021) funded within the Programma Operativo Nazionale (PON) Ricerca e Innovazione of the Italian Ministry of University and Research. Dr. Delrio acknowledges funding from Eurecom.

References

1. V. Randazzo, E. Puleo, A. Paviglianiti, A. Vallan, and E. Pasero, "Development and validation of an algorithm for the digitization of ecg paper images," *Sensors*, vol. 22, no. 19, p. 7138, 2022.
2. G. Cirrincione, V. Randazzo, and E. Pasero, "The growing curvilinear component analysis (gcca) neural network," *Neural Networks*, vol. 103, pp. 108–117, 2018.
3. F. Alfieri, A. Ancona, G. Tripepi, V. Randazzo, A. Paviglianiti, E. Pasero, L. Vecchi, C. Politi, V. Cauda, and R. M. Fagugli, "External validation of a deep-learning model to predict severe acute kidney injury based on urine output changes in critically ill patients," *Journal of Nephrology*, vol. 35, no. 8, pp. 2047–2056, 2022.
4. M. Stefanini, M. Lovino, R. Cucchiara, and E. Ficarra, "Predicting gene and protein expression levels from dna and protein sequences with perceiver," *Computer Methods and Programs in Biomedicine*, vol. 234, p. 107504, 2023.
5. G. Cirrincione, S. Cannata, G. Cicceri, F. Prinzi, T. Currieri, M. Lovino, C. Militello, E. Pasero, and S. Vitabile, "Transformer-based approach to melanoma detection," *Sensors*, vol. 23, no. 12, p. 5677, 2023.

6. E. Pasero, F. Gaita, V. Randazzo, P. Meynet, S. Cannata, P. Maury, and C. Giustetto, "Artificial intelligence ecg analysis in patients with short qt syndrome to predict life-threatening arrhythmic events," *Sensors*, vol. 23, no. 21, p. 8900, 2023.
7. V. Randazzo, G. Marchetti, C. Giustetto, E. Gugliermina, R. Kumar, G. Cirrincione, F. Gaita, and E. Pasero, "Learning-based approach to predict fatal events in brugada syndrome," in *Applications of Artificial Intelligence and Neural Systems to Data Science*, pp. 63–72, Springer, 2023.
8. P. Barbiero, M. Lovino, M. Siviero, G. Ciravegna, V. Randazzo, E. Ficarra, and G. Cirrincione, "Unsupervised multi-omic data fusion: The neural graph learning network," in *Intelligent Computing Theories and Application: 16th International Conference, ICIC 2020, Bari, Italy, October 2–5, 2020, Proceedings, Part I*, pp. 172–182, Springer, 2020.
9. M. Lovino, V. Randazzo, G. Ciravegna, P. Barbiero, E. Ficarra, and G. Cirrincione, "A survey on data integration for multi-omics sample clustering," *Neurocomputing*, vol. 488, pp. 494–508, 2022.
10. M. Lovino, G. Bontempo, G. Cirrincione, and E. Ficarra, "Multi-omics classification on kidney samples exploiting uncertainty-aware models," in *International Conference on Intelligent Computing*, pp. 32–42, Springer, 2020.
11. J. Ferretti, P. Barbiero, V. Randazzo, G. Cirrincione, and E. Pasero, "Towards uncovering feature extraction from temporal signals in deep cnn: the ecg case study," in *2020 international joint conference on neural networks (IJCNN)*, pp. 1–7, IEEE, 2020.
12. F. Delrio, V. Randazzo, G. Cirrincione, and E. Pasero, "Non-invasive arterial blood pressure estimation from electrocardiogram and photoplethysmography signals using a conv1d-bilstm neural network," *Engineering Proceedings*, vol. 39, no. 1, p. 78, 2023.
13. A. Paviglianiti, V. Randazzo, G. Cirrincione, and E. Pasero, "Neural recurrent approches to noninvasive blood pressure estimation," in *2020 international joint conference on neural networks (IJCNN)*, pp. 1–7, IEEE, 2020.
14. A. Paviglianiti, V. Randazzo, E. Pasero, and A. Vallan, "Noninvasive arterial blood pressure estimation using abpnet and vital-ecg," in *2020 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, pp. 1–5, IEEE, 2020.
15. J. Ferretti, V. Randazzo, G. Cirrincione, and E. Pasero, "1-D convolutional neural network for ECG arrhythmia classification," in *Progresses in Artificial Intelligence and Neural Systems* (A. Esposito, M. Faundez-Zanuy, F. C. Morabito, and E. Pasero, eds.), Smart Innovation, Systems and Technologies, pp. 269–279, Springer.
16. G. Cirrincione, V. Randazzo, and E. Pasero, "A neural based comparative analysis for feature extraction from ECG signals," in *Neural Approaches to Dynamics of Signal Exchanges* (A. Esposito, M. Faundez-Zanuy, F. C. Morabito, and E. Pasero, eds.), Smart Innovation, Systems and Technologies, pp. 247–256, Springer.
17. P. Laguna, R. G. Mark, A. Goldberg, and G. B. Moody, "A database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG," in *Computers in cardiology 1997*, pp. 673–676, IEEE, 1997.
18. A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals," *circulation*, vol. 101, no. 23, pp. e215–e220, 2000.
19. "The WFDB software package."
20. L. Debnath and J.-P. Antoine, "Wavelet transforms and their applications," *Physics Today - PHYS TODAY*, vol. 56, pp. 68–68, 04 2003.
21. L. Aguiar-Conraria and M. J. Soares, "The continuous wavelet transform: Moving beyond uni- and bivariate analysis," *Journal of Economic Surveys*, vol. 28, no. 2, pp. 344–375, 2014.
22. X. Liang, L. Li, Y. Liu, D. Chen, X. Wang, S. Hu, J. Wang, H. Zhang, C. Sun, and C. Liu, "ECG_segnet: an ECG delineation model based on the encoder-decoder structure," *Computers in Biology and Medicine*, vol. 145, p. 105445, 2022.
23. N. Hughes, L. Tarassenko, and S. J. Roberts, "Markov models for automated ecg interval analysis," *Advances in Neural Information Processing Systems*, vol. 16, 2003.