Neural Network-based ECG Delineation

Idriss Ngomseu Tchoupe[1], Mously Dior Diaw[1,2], Stéphane Papelier[2], Alexandre Durand-Salmon[2], Julien Oster[1,3]

[1]IADI, U1254, Inserm, Université de Lorraine, Nancy, France
[2]Cardiabase, Banook Group, Nancy, France
[3]CIC-IT 1433, Université de Lorraine, Inserm, CHRU de Nancy, Nancy, France

Abstract

ECG delineation is crucial for assessing drug-induced proarrhythmic risks, but still heavily depends on expert cardiologists despite existing automated methods. Our study proposes a deep learning model, E-DelUnet, to accurately locate key ECG fiducial points: P_{onset}, QRS_{onset}, QRS_{offset}, T_{peak}, and T_{end}.

E-DelUnet is an adapted U-Net without skip connections and up-sampling, improving speed, performance, and reducing overfitting. It processes 1.2s single-lead ECGs and outputs binary masks marking fiducials. We trained the model on 2,054 ECGs from Verapamil and Quinidine studies using 5-fold cross-validation and tested it on 13,085 ECGs from other drug studies.

Compared to traditional U-Net, wavelet-based methods, and residual networks, E-DelUnet showed the best performance, closely matching cardiologist measurements (from 3.86 ± 2.9 ms for the QRS_{onset} to 6.77 ± 7.41 ms for the T_{peak}). This model holds promise for improving ECG analysis in drug safety assessments and clinical workflows.

1. Introduction

Electrocardiography (ECG) plays a crucial role in cardiac diagnostics, offering a non-invasive method to assess the heart electrical activity. Accurate delineation of ECG signals —identifying the exact onset, peak, and offset of the various waveforms (P wave, QRS complex, and T wave)— is vital for diagnosing various cardiac conditions. Traditionally, this task is performed manually by expert cardiologists, a process that is not only time-consuming but also subject to inter-observer variability. While numerous algorithms have been developed to automate this process, including wavelet-based [1] methods and deep learning approaches [2],[3], [4], challenges persist in achieving high accuracy across diverse datasets.

In this study, we propose E-DelUnet, a novel deep learning approach using an adapted U-Net architecture [5] —originally intended for biomedical image segmentation— for ECG signal analysis. E-DelUnet, with eight convolutional layers and a fully connected layer, omits skip connections and up-sampling to meet ECG delineation challenges. This design allows E-DelUnet to accurately identify and localize various ECG fiducial points, such as P_{onset}, QRS_{onset}, QRS_{offset}, T_{peak} and T_{end}, by generating a mask from which these points are extracted. Furthermore, our deep learning model requires minimal parameter tuning and preprocessing, enhancing its ability to generalize effectively to new datasets.

2. Materials and methods

2.1. Dataset and pre-processing

This study utilized three PhysioNet [6] ECG databases: ECGRDVQ [7] (Study 1), ECGDMMLD [8] (Study 2), and ECGCIPA [9] (Study 3), as summarized in Table 1. Semi-automated ECG fiducial point annotations (cardiologist measurements) were reviewed manually by the two independent human readers. For pre-processing, 1.2-second average heartbeats were generated from each 10-second single-lead recording, with the QRS_{peak} centered at 400 ms, resampled to 500 Hz, and normalized (zero mean and unit standard deviation).

Table 1. Data summary. R, D, V, Q, and P are respectively : Ranolazine, Dofetilide, Verapamil, Quinidine and Placebo

Dataset	Nb of Drugs	Nb of Patients	Nb 12-lead ECG recordings
Cross-validation			
Study 1 (Q,V)	2	22	2054
Final Test			
Study 1 (D,R,P)	2	22	3162
study 2	4	22	4195
study 3	4	60	5728

2.2. Problem formulation and ECG template

We aim to detect the start of the P wave, the start and end of the QRS complex, and the peak and end of the T wave, referred to as P_{onset}, QRS_{onset}, QRS_{offset}, T_{peak} and T_{end}, respectively. We approach this as a segmentation task, where the model inputs the ECG heartbeat and outputs a binary mask matching the ECG size. Each transition from 0 to 1 or 1 to 0 in the mask corresponds to a fiducial point of interest, as shown in Figure 1.B.

2.3. The proposed model

Inspired by the U-Net architecture [5], originally designed for biomedical image segmentation, we developed E-DelUnet specifically for ECG delineation, as shown in Figure 1.A. Our model omits skip connections and up-sampling layers to reduce overfitting and complexity. Instead of up-sampling, a fully connected layer restores the input signal size, as multiple up-samplings might not add significant value if the network has already learned the relevant information.

E-DelUnet has 2,447,352 parameters and consists of three parts. Part 1 includes four 1D convolutional layers (kernel size = 7, padding = 3) with increasing numbers of filters (64, 128, 256, 512), each followed by ReLU, Batch Normalization, dropout (p=0.41), and 2x1 max pooling. Part 2 features four 1D convolutional layers (kernel size = 7, padding = 3) with decreasing numbers of filters (256, 128, 64, 32), maintaining the signal size at 37 samples, also followed by ReLU, Batch Normalization, and dropout (p=0.41). Part 3 consists of adaptive average pooling, a fully connected layer, and a sigmoid function for output.

This architecture captures intricate features in Part 1 through increasing filters and down-sampling, refines these features in Part 2 while maintaining signal size, and compresses them in Part 3 to produce a probabilistic output.

2.4. Details of training and evaluation process

The three datasets in Table 1 were divided into training, validation, and test sets. Quinidine (Q) and Verapamil (V) ECG recordings from Study 1, totaling 2,054 recordings, were used for training and validation, while the remaining data (Study 1 (D, R, P), Study 2, and Study 3) were reserved for testing. We employed patient-stratified 5-fold cross-validation on the training dataset, ensuring data from the same patient remained in the same fold. Final test results were obtained by averaging the fiducial point positions extracted from the masks produced by each of the five models.

The model was trained using Binary Cross-Entropy (BCE) as the loss function and assessed using mean absolute error (MAE) of the time errors to measure localization accuracy. Training was conducted over 60 epochs with a batch size of 32. The best model for each fold was saved based on the lowest validation BCE. Weights were initialized using the Xavier initializer, and the Adam optimizer was applied with a learning rate of 0.0032. The learning rate and dropout probability were optimized using the Optuna algorithm [10]

3. Results and comparisons

The final results were obtained by averaging positions from five models generated through 5-fold cross-validation. We combined the twelve masks from each lead of a 12-lead ECG into a single mask, following the same approach used by cardiologists when annotating the datasets. Fiducial point values were converted from samples to milliseconds.

Table 2 shows our model performance (MAE ± std) across the different test sets, with MAEs ranging from 3.86 ± 2.9 ms for QRS_{onset} to 6.77 ± 7.41 ms for T_{peak}. E-DelUnet outperformed other algorithms (ResNet, base U-Net, and a wavelet-based method) on most fiducial points.

Figure 2.a highlights an example of the model's prediction. Predicted (red) and reference (blue) mask, and ECG heartbeat (orange). Figures 2.b to 2.f show Bland-Altman plots with biases and 95% limits of agreement (LOA) and scatter plots comparing predictions to ground truth. The scatter plots indicate strong correlations for P_{onset}, T_{peak}, and $Tend$, with more spread for QRS_{onset} and QRS_{offset}. Overall, Bland-Altman plots confirm that our model's automatic measurements align well with manual ones, with acceptable biases and variability.

4. Discussion and Conclusion

This study introduces E-DelUnet, a deep learning model designed for automated ECG delineation, crucial for diagnosing cardiac conditions and assessing drug-induced cardiotoxicity. Utilizing an adapted U-Net architecture and trained on extensive PhysioNet datasets, E-DelUnet demonstrates superior performance compared to U-Net, ResNet, and wavelet-based models we tested. It effectively identifies key ECG fiducial points, including P_{onset}, QRS_{onset}, QRS_{offset}, T_{peak}, and T_{end}.

The model has shown good performance across diverse datasets: Study 1 (D, R, P) (ECGRDVQ), 2(ECGDMMLD) and 3(CiPA), indicating good generalization capabilities. However, challenges persist with accurately identifying QRS_{onset} and QRS_{offset} due to limited reference values. This is made evidently clear by the BlandAltman depiction, where the proposed model seem to be only guessing a position in a given range. Bland Altman depiction seems also to be highlighting inaccura-

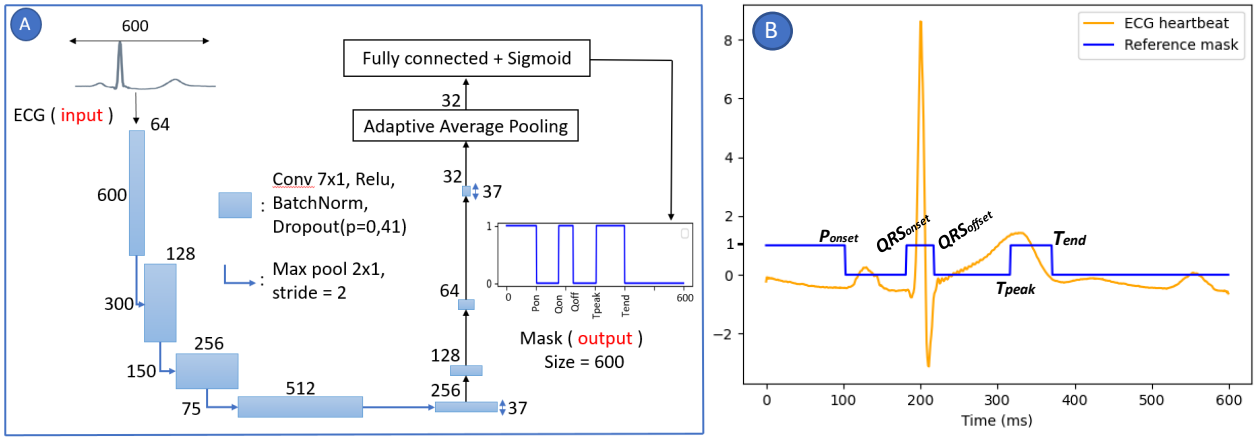

Figure 1. Figure **A** shows the E-DelUnet architecture, which takes a 600-sample ECG signal as input and generates an output mask of the same length. This mask indicates the locations of each fiducial point of interest. Figure **B** displays the ECG heartbeat (orange) alongside its corresponding mask (blue).

Table 2. Comparisons of E-DelUnet with others methods for automatic ECG delineation (MAE ± STD). **Bold** is the best.

Test data	Methods	P_{onset}	QRS_{onset}	QRS_{offset}	T_{peak}	T_{end}	Parameters
ECGRDVQ (R,D,P)	U-Net	4.68 ± 4.02	2.77 ± 2.1	**5.44 ± 4.31**	4.85 ± 5.71	4.99 ± 6.68	10,455,489
	Wavelet	25.82 ± 16.98	24.4 ± 7.6	10.93 ± 7.39	15.38 ± 12.2	17.23 ± 16.12	.....
	ResNet	4.58 ± 4.26	**2.35 ± 1.95**	5.62 ± 4.23	6.28 ± 8.02	7.12 ± 9.1	1,656,066
	E-DelUnet	**4.2 ± 3.79**	2.89 ± 2.12	5.45 ± 4.63	**4.31 ± 5.44**	**4.31 ± 6.64**	2,447,352
ECGDMMLD	U-Net	6.53 ± 4.71	**2.86 ± 2.03**	9.81 ± 6.09	**6.91 ± 6.51**	**5.95 ± 4.47**	
	Wavelet	33.84 ± 18.46	26.62 ± 8.26	18.81 ± 10.75	22.61 ± 15.26	15.08 ± 12.15	
	ResNet	6.6 ± 5.75	2.98 ± 1.97	10.83 ± 6.05	10.79 ± 9.8	6.37 ± 6.6	
	E-DelUnet	**5.91 ± 4.48**	3.22 ± 2.07	**8.52 ± 5.83**	6.98 ± 7.84	6.05 ± 4.48	
CiPA	U-Net	5.13 ± 4.97	4.71 ± 3.32	5.72 ± 4.51	**7.49 ± 7.88**	6.83 ± 7.11	
	Wavelet	31.7 ± 19.5	20.94 ± 7.23	12.49 ± 8.32	17.66 ± 15.23	17.3 ± 13.8	
	ResNet	5.99 ± 7.12	**4.19 ± 2.93**	6.53 ± 4.94	9.0 ± 10.19	10.18 ± 9.13	
	E-DelUnet	**4.69 ± 4.47**	4.86 ± 3.42	**5.44 ± 3.99**	7.97 ± 7.7	**6.75 ± 6.97**	
Global test set	U-Net	5.47 ± 4.73	3.65 ± 2.84	6.96 ± 5.4	6.80 ± 7.05	6.1 ± 6.31	
	Wavelet	30.96 ± 18.83	23.6 ± 8.05	14.14 ± 9.55	18.7 ± 14.84	16.57 ± 13.95	
	ResNet	5.84 ± 6.14	**3.36 ± 2.55**	7.69 ± 5.62	8.91 ± 9.73	8.22 ± 8.58	
	E-DelUnet	**4.96 ± 4.37**	3.86 ± 2.9	**6.43 ± 5.01**	**6.77 ± 7.41**	**5.93 ± 6.27**	

cies in the positioning of T_{peak} and T_{end} on complex T wave morphologies. Efforts to improve these areas with weighted masks were not successful, suggesting a need for a more varied training dataset.

Overall, E-DelUnet shows promise in reducing the reliance on manual annotation by cardiologists and improving cardiac assessments, with future research aimed at refining the model, expanding the dataset, and validating its clinical utility.

Acknowledgments

This study was funded thanks to a Grand Est Research Development and Innovation project (AFICIONADO).

References

[1] Martínez JP, et al. A wavelet-based ecg delineator: Evaluation on standard databases. IEEE TBME 2004;51.

[2] Chen L, et al. Ecgvednet: A variational encoder-decoder network for ecg delineation in morphology variant ecgs. IEEE TBME 2024;.

[3] Jimenez-Perez G, et al. Ecg-delnet: Delineation of ambulatory electrocardiograms with mixed quality labeling using neural networks. ArXiv Preprint 2020;.

[4] Peimankar A, et al. Dens-ecg: A deep learning approach for ecg signal delineation. ESA 2021;165.

[5] Ronneberger O, et al. U-net: Convolutional networks for biomedical image segmentation. ArXiv 2015;.

[6] Goldberger AL, et al. Physiobank, physiotoolkit, and phys-

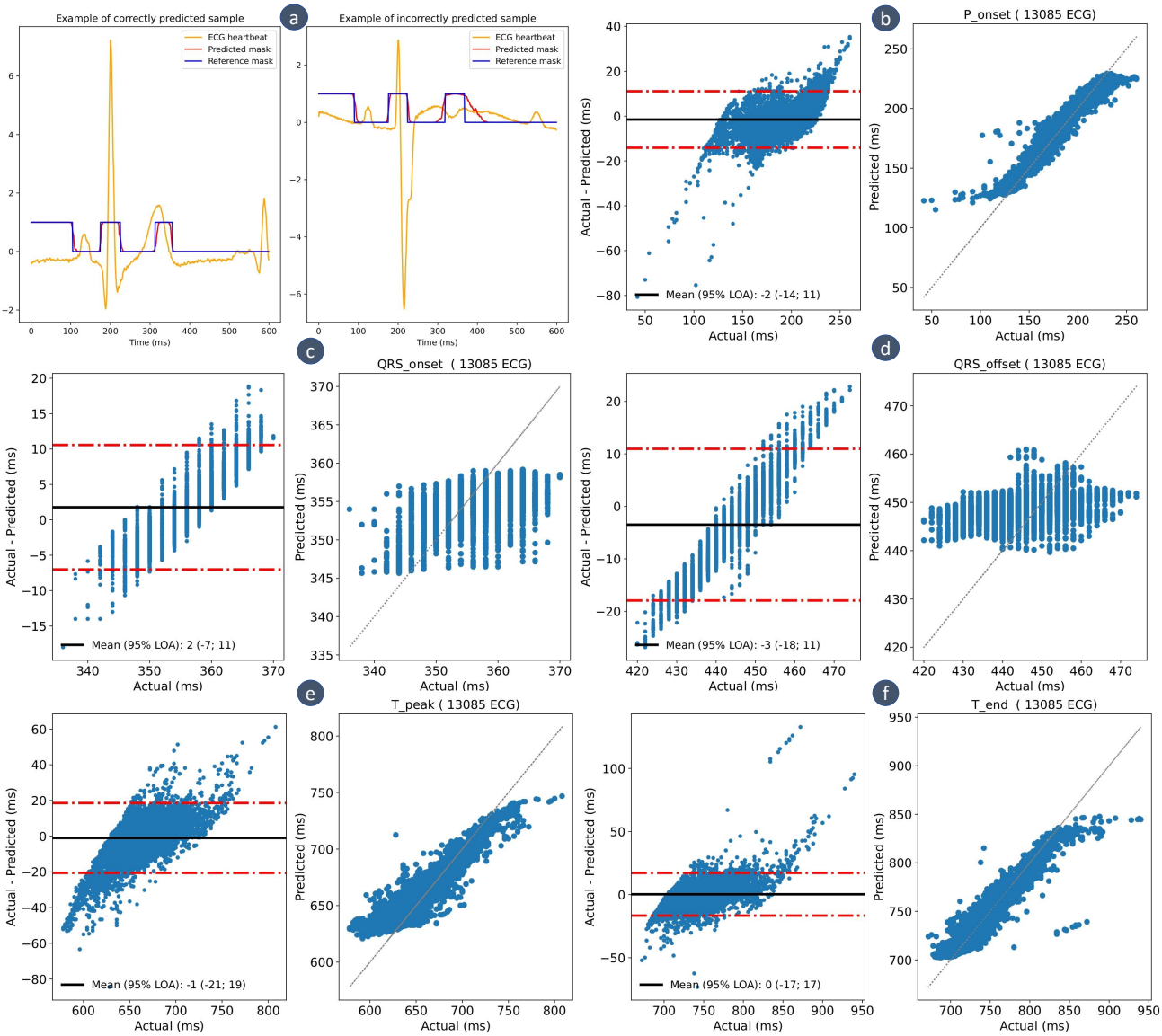

Figure 2. (a) Sample visualization and (b→f) Bland-Altmann (left) and scatterplots (right) of manual versus automatic measurements of P_{onset}, QRS_{onset}, QRS_{offset}, T_{peak} and T_{end}. The black line on the Bland-Altmann plots corresponds to the mean difference and the red lines to the limits of agreement (LOA: mean ± 1.96 std). The black line on the scatterplots corresponds to an ideal fit where manual and automatic delineation are in perfect agreement (y=x). The blue markers represent data points from Study 1 (D, R, P), 2 and 3.

ionet: components of a new research resource for complex physiologic signals. Circulation 2000;101.

[7] Johannesen L, et al. Differentiating drug-induced multichannel block on the electrocardiogram: randomized study of dofetilide, quinidine, ranolazine, and verapamil. Clin Pharmacology Therapeutics 2014;96.

[8] Johannesen L, et al. Late sodium current block for drug-induced long qt syndrome: results from a prospective clinical trial. Clin Pharma Therap 2016;99.

[9] Vicente J, et al. Assessment of multi-ion channel block in a phase i randomized study design: Results of the ci-pa phase i ecg biomarker validation study. Clin Pharm Thera 2019;.

[10] Akiba T, et al. Optuna: A next-generation hyperparameter optimization framework. In Proceedings of the 25th ACM SIGKDD Intern Conf on KDD. 2019; 2623–2631.

Address for correspondence:

Julien Oster (julien.oster@inserm.fr)
Batiment Recherche, Rue du Morvan, 54511 Vandœuvre-les-Nancy, France