Layout-Invariant U-Net Segmentation for Time-Series Reconstruction from Noisy ECG Images

Samer Jammoul[1], Abdullatif Hassan[1], Emily Zhang[1], Philip A Warrick[1,2], Jonathan Afilalo[1]

[1]McGill University, Montreal, Canada, [2]PeriGen Inc., Montreal, Canada

Abstract

ECGs are routinely stored and exchanged as images rather than time-series. Converting ECG images to time-series waveforms would render them accessible to models relying on time-series format, and enhancing their utility at the bedside. We developed a novel automated pipeline to extract the underlying time-series waveforms from images of standard 12-lead ECGs. Framed as an image segmentation task, our algorithm consists of image de-skewing and reshaping, resolution estimation, and signal isolation to remove noise, distortion and gridlines. Finally, we convert the pixel coordinates of the isolated signal on the image into a time-series waveform and assign it to the correct lead. For the segmentation task, we trained a U-Net neural network model using synthetically generated images of standard 12-lead ECGs. To render the model robust to noise and distortions, we augmented the images with random distortions such as handwriting, paper wrinkles, image cropping and noise. For training, we used the 21,799 records of the PTB-XL ECG database. For the digitization task, our team "ECGénie" obtained an official score on the hidden test data of 5.037 SNR (signal-to-noise ratio). In local five-fold cross-validation, the SNR for the validation data was 10.332±0.465 (mean±std).

1. Introduction

Many models have been developed to analyze ECG time-series signals to predict parameters of cardiac function and arrhythmias (1; 2; 3; 4). However, a significant portion of ECG data is not readily available in time-series format. Often, at the bedside, only the physical paper ECG is accessible to healthcare workers, and they may communicate by exchanging images of these ECGs. Additionally, some hospitals store ECGs in electronic health records as scanned images, rather than in a more analyzable digital format.

Neural-network-based models developed for ECG-based diagnosis of arrhythmias are now comparable to experts in terms of accuracy (5). As well, there are models for ECG interpretations that are impossible for experts, such as such as sex, age or left-ventricle ejection fraction. In this context, conversion of ECG images to time-series could provide more flexibility to the automated ECG diagnosis workflow.

Therefore, the objective of this work is to develop an automated pipeline to convert an ECG image to a digital time-series signal. Perhaps surprisingly, this problem has scant presence in the literature; the George B. Moody PhysioNet Challenge 2024 addresses the potential to develop this research area. The Challenge description (6) cites five references from 2012 to 2021: the earlier works used image processing alone, while two used deep-learning approaches such as multilayer perceptrons (MLPs) or U-Net (7; 8). Two approaches in related papers from Computing in Cardiology 2023 used a random forest classifier for semantic segmentation (9) or an architecture with ResNet50 building blocks (10).

As a contribution to the Challenge, we used the U-Net architecture (11) for the main segmentation task in this work, followed by post-processing to obtain time-series. The development of the U-Net deep-neural-network architecture was initially motivated by the need for improved image segmentation of electron and light microscopy images. Since then it has been used successfully in many segmentation tasks within and beyond the biomedical field. The novelty of our approach includes ECG layout invariance through the use of a row-based targets, and our provision of numerous extensions to the standard augmentations provided by the ECG Image Kit.

2. Methods

2.1. Data

The PhysioNet Challenge 2024 provided ECG data in the form of time series and photos (12) as well as software tools to generate images via the ECG Image Kit (13; 14). The 21,799 12-channel ECG records of the PTB-XL database (15; 16) was the main source of time-series data. The associated PTB-XL+ dataset provides ECG features generated from three algorithms: GE Healthcare Mar-

 ISSN: 2325-887X DOI: 10.22489/CinC.2024.481

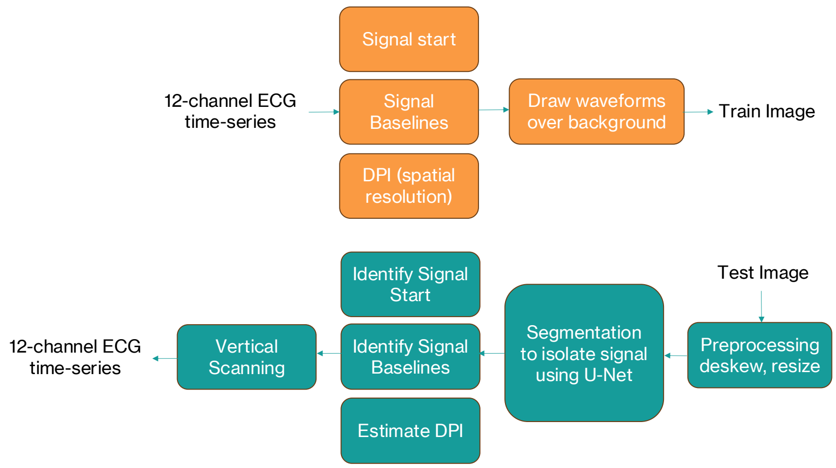

Figure 1. Overall approach. Image generation (top). Image segmentation and signal reconstruction (bottom).

quette 12SL (17), University of Glasgow (18) (both commercial), and the open-source algorithm ECGDeli (19). It also provides ECG class-mappings to standard nomenclatures such as SNOMED CT (20) and LOINC (21). These provided ground-truth labels for the Challenge classification task, which we omitted in this work.

2.2. Overall Approach

Fig. 1 shows a block diagram of the two symmetric ECG processing steps: 1) image generation from time-series for training data and 2) time-series from segmentation and post-processing of images.

2.3. Image generation from time-series

For the Challenge digitization task, we used and modified the ECG Image Kit to obtain labelled, synthetically generated images with numerous added noise and distortions. In addition to the transformations provided by the ECG Image Kit, we added randomization of fonts, label translations, signal translation, layout spacing, pulse calibration morphologies, grid patterns and cropping. For further augmentation, we used the augraphy python package to emulate common noise seen in paper documents, including bad printer, faxing, bad photocopy and shadows.

We considered all common layouts that follow the standard pattern of 3 top rows of 4 lead segments of 2.5 s duration having the following fixed row and lead ordering: I, aVR, V1, V4; II, aVL, V2, V5; III, aVF, V3, V6. We anticipated variants on this pattern by randomization of the number (from 0-3) and selection of the subsequent rows of rhythm leads. See Figure 2.

In addition to these variants for training augmentation, we generated lead-contour images to provide a ground-truth for the subsequent segmentation step. These images labelled the pixels of the signal contours corresponding to their row (from 1-7); otherwise background pixels were labelled 0.

For each of the 21,799 PTB-XL ECGs, we generated five augmentations for a total of 108,995 synthetic training/ground-truth image pairs. We chose this as a reasonable tradeoff between the benefits of augmentation vs. training time.

2.4. Time-series from images

Starting from an arbitrary input image, either synthetic from the image generation or real from photos of paper ECGs, we first deskewed the image to remove global rotations and then resize the width and height to the nearest power of 2. These images were then presented to a U-Net architecture using binned batching to accommodate the variable image sizes. We trained for a total of 5 epochs using a learning rate of 10^{-5}. We used data partitions of 80% training and 20% testing for each fold of 5-fold cross-validation. The validation set, 10% of training, was used to assess generalization status.

At inference time, we applied several post-processing steps to the resulting segmentation to obtain the time-series. First, we identified the bounding box of the layout rows by searching from each of four edges for the first transition of the pixel row or column summation that was above an empirically chosen threshold. From the maximum output segmentation value we inferred the number of layout rows (4-7), the bounding boxes of each row and corresponding DPI. For each layout row, we identified the signal baselines from the maximum pixel row summation.

Finally, for each pixel column, we inferred a time based on the DPI and an amplitude value in μV from the maximum pixel rows. We chose rhythm leads from the corresponding 2.5 s lead segment(s) with the highest correlation.

3. Results

Figure 3 shows an example test output segmentation image. The segmentation performance over 5 folds was an AUPRC of 0.9334±0.0122, a precision of 0.8905±0.0197 and a sensitivity of 0.8990±0.0162.

To evaluate the fidelity of the extracted time-series waveforms, we used the Challenge code to compare the extracted waveforms to the ground-truth waveforms and compute the average signal-to-noise ratio (SNR) of all 12 leads expressed on a logarithmic decibel scale. The signal-to-noise ratios (SNR) of our estimated time series for different data types are shown in Table 1. To summarize, for the digitization task our team "ECGénie" obtained official and hackathon overall SNR scores on the hidden test data of 5.037 and 0.057, respectively. In local five-fold cross-validation, the SNR for the validation data was 10.332±0.465 (mean±std).

Table 1. Digitization task: overall and component scores

Data Type	**SNR$_O^*$**	**SNR$_H^*$**
Overall score (unranked)	5.037	0.057
Color scans of clean papers	5.182	0.054
Black-and-white scans of clean papers	4.89	0.060
Mobile phone photos of clean papers	-4.936	-3.017
Mobile phone photos of stained papers	-5.746	-3.080
Mobile phone photos of deteriorated papers	-1.975	-1.771
Color scans of deteriorated papers	-0.173	-0.193
Black-and-white scans of deteriorated papers	-0.228	-0.189
Screenshots of computer monitor	-7.771	-3.787

*SNR$_O$: official results; SNR$_H$: hackathon results

4. Discussion

These are encouraging preliminary results. We will evaluate the model's performance on photographs of real ECGs and the reliability of the reconstructed time-series data for subsequent analysis and interpretation. We found that even when the signal segmentation was excellent, signal fidelity was very sensitive to errors in DPI estimation, deskewing and perspective correction. In future work, we would like to demonstrate that the resulting extracted signal is on par with the raw signal for use in clinical practice.

References

[1] Gupta U, Paluru N, Nankani D, Kulkarni K, Awasthi N. A Comprehensive Review on Efficient Artificial Intelligence Models for Classification of Abnormal Cardiac Rhythms using Electrocardiograms. Heliyon March 2024;10(5):e26787. ISSN 2405-8440. Place: England.

[2] Ansari Y, Mourad O, Qaraqe K, Serpedin E. Deep learning for ECG Arrhythmia Detection and Classification: an Overview of Progress for Period 2017-2023. Frontiers in Physiology 2023;14.

[3] Reyna MA, Sadr N, Alday EAP, Gu A, Shah AJ, Robichaux C, Rad AB, Elola A, Seyedi S, Ansari S, Ghanbari H, Li Q, Sharma A, Clifford GD. Issues in the Automated Classification of Multilead ECGs using Heterogeneous Labels and Populations. Physiological Measurement 2022;.

[4] Alday EAP, Gu A, Shah AJ, Robichaux C, Wong AKI, Liu C, Liu F, Rad AB, Elola A, Seyedi S, Li Q, Sharma A, Clifford GD, Reyna MA. Classification of 12-Lead ECGs: the PhysioNet/Computing in Cardiology Challenge 2020. Physiological Measurement January 2021;41(12).

[5] Kashou AH, Mulpuru SK, Deshmukh AJ, Ko WY, Attia ZI, Carter RE, Friedman PA, Noseworthy PA. An Artificial Intelligence–enabled ECG Algorithm for Comprehensive ECG Interpretation: Can it Pass the 'Turing Test'? Cardiovascular Digital Health Journal June 2021;2(3):164–170. ISSN 2666-6936.

[6] Reyna MA, Deepanshi, Weigle J, Koscova Z, Elola A, Seyedi S, Campbell K, Clifford GD, Sameni R. Digitization and Classification of ECG Images: The George B. Moody PhysioNet Challenge 2024. Computing in Cardiology 2024;51:1–4.

[7] Mishra S, Khatwani G, Patil R, Sapariya D, Shah V, Parmar D, Dinesh S, Daphal P, Mehendale N. ECG Paper Record Digitization and Diagnosis Using Deep Learning. Journal of Medical and Biological Engineering August 2021;41(4):422–432. ISSN 2199-4757.

[8] Li Y, Qu Q, Wang M, Yu L, Wang J, Shen L, He K. Deep Learning for Digitizing Highly Noisy Paper-based ECG Records. Computers in Biology and Medicine December 2020;127:104077. ISSN 1879-0534 0010-4825.

[9] Melo D, Madeiro1 J, Rigo L, Pessoa1 C, Macêdo1 J, D. G. A Semantic Segmentation-based Digitization of ECG Papers. Computing in Cardiology 2023; 51:1–4.

[10] Hernández-Fernández C, Gómez-Moreno H, R HC,

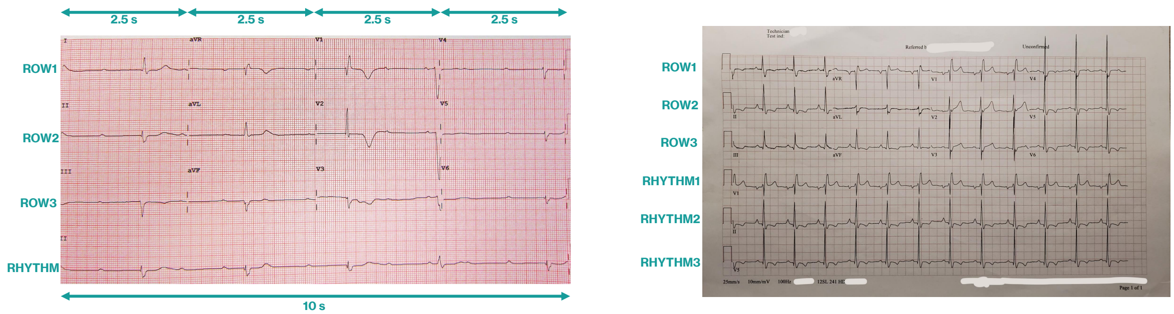

Figure 2. Multilayout support: two example layouts supported by our per-row target approach. Both have 3 rows of four 2.5 s lead segments, followed by the rhythm leads: 1 row (left) or 3 rows (right).

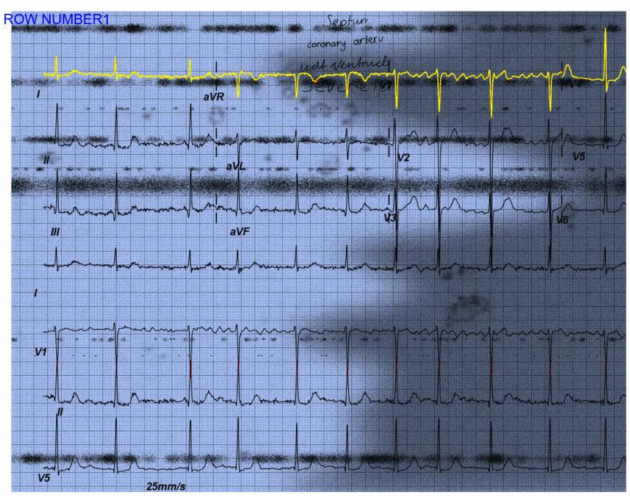

Figure 3. Example test results from a generated image with multiple distortions (annotations, blotches, etc.). The segmentation of the first row is shown in yellow for the 2.5 s segments from leads I, II, V1, and V4.

Blanco-Velasco M. Long-term ECG Analysis Through Image Conversion and Deep Learning. Computing in Cardiology 2023;51:1–4.

[11] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. CoRR 2015;abs/1505.04597.

[12] Reyna MA, Deepanshi WJ, Z K, K C, KK S, Saghafi S, Nikookar S, M MS, Kiarashi Y, Seyedi S, Clifford G, R. S. ECG-Image-Database: A Dataset of ECG Images with Real-world Imaging and Scanning Artifacts; a Foundation for Computerized ECG Image Digitization and Analysis. Arxiv 2024;URL https://arxiv.org/abs/2409.16612.

[13] Shivashankara KK, Deepanshi, Shervedani AM, Reyna MA, Clifford GD, Sameni R. ECG-Image-Kit: a Synthetic Image Generation Toolbox to Facilitate Deep Learning-based Electrocardiogram Digitization. Physiological Measurement 2024;45:055019.

[14] Deepanshi, Shivashankara KK, Clifford GD, Reyna MA, Sameni R. ECG-Image-Kit: A Toolkit for Synthesis, Analysis, and Digitization of Electrocardiogram Images, January 2024. Online at: https://github.com/alphanumericslab/ecg-image-kit.

[15] Wagner P, Strodthoff N, Bousseljot RD, Kreiseler D, Lunze FI, Samek W, Schaeffter T. PTB-XL, a Large Publicly Available Electrocardiography Dataset. Scientific Data 2020;7:154.

[16] Strodthoff N, Mehari T, Nagel C, Aston PJ, Sundar A, Graff C, Kanters JK, Haverkamp W, Dössel O, Loewe A, Bär M, Schaeffter T. PTB-XL+, a Comprehensive Electrocardiographic Feature Dataset. Scientific Data 2023;10:279.

[17] GE Healthcare. Marquette 12SL ECG Analysis Program: Physician's Guide, version 2056246-002c. Scientific Data 2019;.

[18] Macfarlane P, Devine B, E C. The University of Glasgow (Uni-G) ECG Analysis Program. Computing in Cardiology 2005;51:1–4.

[19] Pilia N, Nagel C, Lenis G, Becker S, Dössel O, Loewe A. ECGdeli - an Open Source ECG Delineation Toolbox for MATLAB. SoftwareX 2021; 13(100639).

[20] National Library of Medicine. Health Information Technology and Health Data Standards at NLM. National Library of Medicine 2022;.

[21] Regenstrief Institute. Logic Observation Identifier Names and Codes. Regenstrief Institute 2022;URL https://loinc.org/.

Address for correspondence: Samer Jammoul, Lady Davis Institute for Medical Research, Jewish General Hospital, 3755 Côte Ste-Catherine Road, Montreal QC, H3T 1E2 Canada, samer.jammoul@mail.mcgill.ca