U-Net Guided Digitization of 12-Lead Printed ECGs

Álvaro José Bocanegra[1], Etel Silva Garcia[2], Andrea Saglietto[3,4], Oscar Camara[1]

[1] Universitat Pompeu Fabra, Barcelona, Spain
[2] Hospital Universitario Puerta del Mar, Cádiz, Spain
[3] University of Turin, Turin, Italy
[4] "Città della Salute e della Scienza" Hospital, Turin, Italy

Abstract

Team Berru's Widows presents a solution for the Digitization of ECG Images in The George B. Moody Physionet Challenge 2024. Although we were not ranked on the official test data, our approach integrates traditional image processing techniques with deep learning methods, which we consider relevant for the state of the art. Our pipeline consisted on a preprocessing stage, then a single lead isolation process and the conversion from image to signal. The preprocessing included shadow removal, using classic computer vision methods, and segmentation using a U-Net. The other stages included single lead isolation and signal digitization.

1. Introduction

The George B. Moody Physionet Challenge 2024 focused on the digitization and classification of electrocardiograms (ECGs) captured from images or paper printouts [1]. This challenge included two tasks: digitization and classification, with the digitization task being the focus of our work. The evaluation was based on comparing the acquired signal against the original digital signal using the signal-to-noise ratio.

The dataset used was the PTB-XL [2], hosted on Physionet [3], which includes 21,799 clinical 12-lead ECG records. Although these data were originally in digital format, the organizers provided the ECG-Image-Kit [4, 5] to generate synthetic ECG images on paper-like backgrounds with various distortions. For the final ranking, the organizers used a larger dataset including several artifacts [6], however, this solution failed to score in this dataset.

Our approach builds on methodologies like those proposed by [7], integrating grid removal, ECG baseline detection, lead isolation, and single-lead extraction. This approach combines Deep Learning and traditional computer vision methods in the preprocessing steps.

Other works have focused on practical applications, The biggest potential usage of this tool focuses on quick diagnosis for non-experts. Lampreave et al. [8] developed a digitization tool for ECGs using a Virtual Reality device. Isabel et al. [9] translated this work and made it work in a mobile device, which is more suitable for quick diagnosis, as the ones done in emergency units or ambulances. These are fast approaches that look for optimization and fast inference for the limited resources, inference in these devices is currently limited due to GPU hardware limitations.

Our approach aims to generate large datasets from available digital records, contributing to the development of advanced ECG analysis tools. The potential extension of our models to practical applications could bridge the gap between large-scale data generation and real-time clinical use, enhancing diagnostic and treatment solutions.

2. Materials and methods

2.1. Grid removal

Grid removal was accomplished using a U-Net neural network architecture [10]. The U-Net model employed a symmetric structure consisting of an encoder (contracting path) and a decoder (expansion path).

In the contracting path, the network consisted of five convolutional blocks, each with two convolutional layers (kernel size 3), followed by Batch Normalization and ReLU activation. Max-pooling layers with a 2x2 kernel and a dropout layer (rate 0.05) followed each block. The expansion path mirrored the contracting path, using transposed convolutional layers to progressively increase the spatial dimensions while reducing the depth. The input image is resized into patches of 128x128x3 to reduce computing time and complexity in the model, an scheme of the architecture used can be seen in figure 1. Patching considered any image size, padding the patch in the borders to fit the required size. The output image is a binary image of 128x128x1 with the signal segmented.

The AdaBound optimizer was used for training, with binary cross-entropy as the loss function and accuracy as the

 ISSN: 2325-887X DOI: 10.22489/CinC.2024.265

evaluation metric.

Figure 1: U-Net architecture used.

For the training, two sets of images from the dataset were generated. Firstly, a noisy set with shadows, textures and handwritten text, an example can be seen in 2. Secondly, a clean gridless dataset, that can be seen in figure 3. The first dataset is used as training set, after preprocessing, and the second set is used as target.

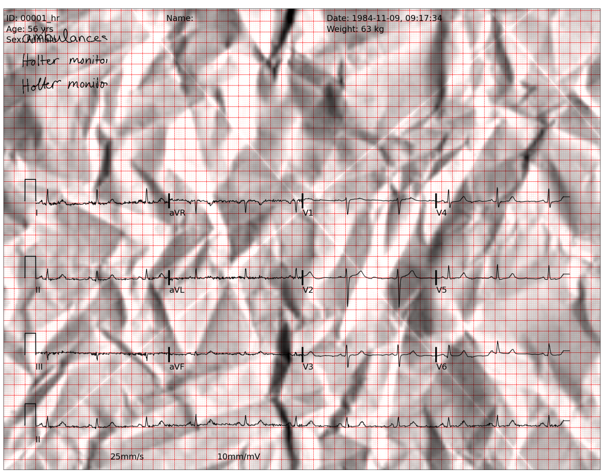

Figure 2: Training image generated with distortions.

For the network to work with a large amount of data and the resources provided by the organizers, the image was resized to a fixed size of 512x384. This size preserved the aspect ratio while reducing the amount of patched in the training process.

2.2. Preprocessing

Preprocessing was performed prior to grid removal to eliminate shadows and normalize intensity. All steps were developed using the OpenCV library [11]. The image was separated into R, G, and B planes, each treated independently. We applied image dilation to preserve edges and a median blur filter to approximate the background, focusing on relevant high-frequency information. By subtracting this background from the original image, we reduced pixel intensity range, isolating high-frequency components. Finally, normalization was applied to utilize the full dynamic

Figure 3: Target image generated without gridlines or any distortion.

range, as shown in Figure 4. The Otsu thresholding method was used to convert the result into a binary image.

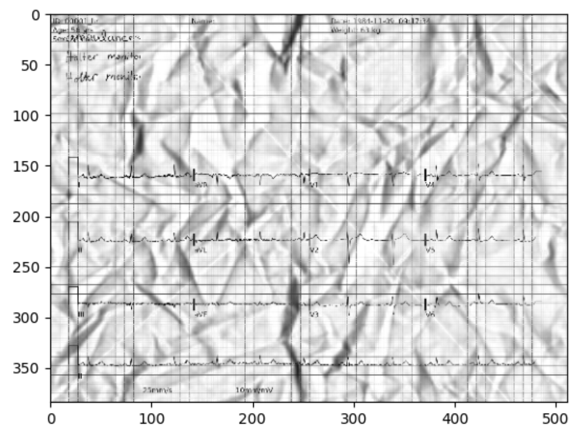

Figure 4: ECG synthetic image with shadows removed.

2.3. Single lead ECG extraction

After shadow and grid removal, we reconstructed the image from the patches. Then, we created an intensity profile of the image rows to identify zones corresponding to ECG baselines (Figure 5). This method assumes no rotation in the image and a standard 3-row by 4-column format.

To isolate each lead, we identified reference points in the image. We assumed that the images were centred equally, and we chose fixed values foe the vertical limits. Then, having the vertical limits and the baselines, we started the isolation process. For extreme limits (e.g., the

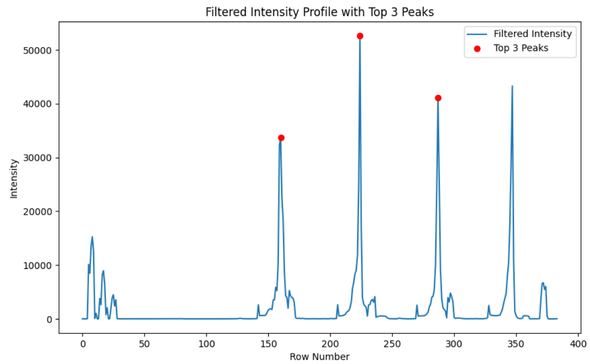

Figure 5: Intensity profile of an image, with red points representing rows likely to be baseline anchors.

upper limit of the first baseline), we calculated the sum of pixel intensities along a vertical line and adjusted until the sum reached zero or reached the end of the image. For inter-baseline limits, we identified the first and last minima within the range. We repeated this process for each baseline space between the vertical limits, this gave us the individual leads with independent heights.

Once isolated, each lead was processed as an individual image, using the average row intensity as the signal baseline, then, the average position of the white pixels per column was used to estimate the value of the signal. Missing values were interpolated using the average of the closest known values, and the signal was resampled to the sampling frequency retrieved from the original files, if no sampling frequency was provided, we assumed 500 Hz. Finally, we scaled based on the image size.

3. Results

The results of our methodology are illustrated in three stages: grid removal, lead segmentation, and signal reconstruction.

Figure 6 shows the output after the grid removal and preprocessing steps, resulting in a clean mask that highlights the ECG signals. Figure 7 shows the segmented leads with red rectangles adapting to their respective ranges. Finally, Figure 8 demonstrates the reconstruction of the ECG signals using the ECG-Image-Kit tool [4, 5].

We evaluated our results using the signal-to-noise ratio (SNR), obtaining a value of -0.849. This value indicates that the noise level was higher than expected, reflecting the challenges of accurately reconstructing ECG signals from images.

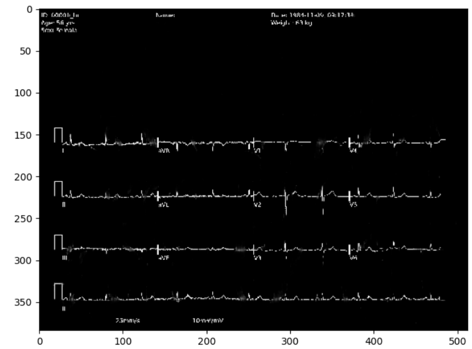

Figure 6: Grid removal and preprocessing results.

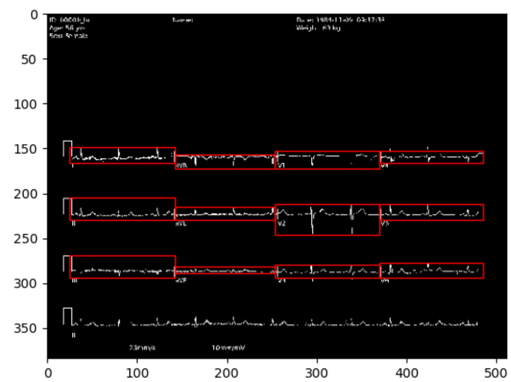

Figure 7: Lead segmentation results. Red rectangles indicate detected and segmented individual leads.

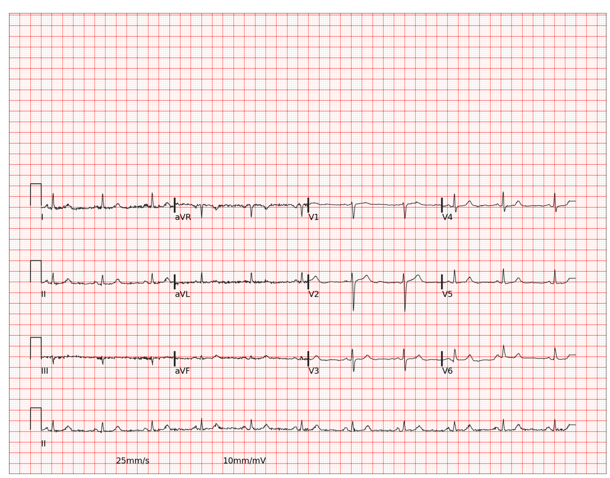

Figure 8: Reconstructed synthetic ECGs generated using the ECG-Image-Kit tool.

4. Discussion and conclusions

The results indicate that our approach, while innovative in integrating deep learning with classical image processing, did not achieve the desired SNR. One key drawback was the image size. Due to the large volume of images processed during the challenge, the original sizes were drastically reduced, lowering patch quality. This led to poor segmentation, errors in single lead isolation, and low resolution in digitization. Potential solutions include reducing the number of images used in training, increasing hardware resources, or reordering the pipeline by performing single lead isolation before segmentation, allowing the use of smaller images in training.

Additionally, our approach made several assumptions regarding the images during the digitization process. The challenge dataset [6] contained rotated images and varying formats, with some ECGs scanned from paper or screens containing extra non-ECG data. To handle such variations, an extended preprocessing step would be necessary, including region of interest detection and rotation correction.

ECG digitization and printed ECG classification pose challenges in real-world applications. While useful for dataset augmentation, printed ECGs suffer from quality loss during digitization, potentially obscuring critical details. Further, lack of temporal alignment between leads in printed ECGs complicates detecting cardiac events that rely on synchronized analysis, such as premature ventricular contractions.

In practical settings, rapid ECG assessments using portable devices, like smartphones, could benefit non-specialists in emergency or pre-hospital environments. These tools could expedite critical decisions where cardiologists are unavailable, potentially improving outcomes. However, current limitations such as image quality, lead synchronization, and preprocessing must be addressed for reliable real-time applications [8, 9].

Smartphones offer a promising platform for rapid ECG digitization but are constrained by limited resources. Our lightweight multi-step approach, which combines U-Net with traditional methods, aligns well with these computational limits, offering a practical solution for resource-constrained applications. Future work should focus on refining these techniques to improve accuracy and efficiency, enhancing their real-world applicability in clinical and emergency settings.

References

[1] 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.

[2] 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(1):1–15.

[3] Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation 2000;101(23):e215–e220.

[4] Shivashankara KK, Shervedani AM, Clifford GD, Reyna MA, Sameni R, et al. ECG-Image-Kit: a synthetic image generation toolbox to facilitate deep learning-based electrocardiogram digitization. Physiological Measurement 2024; 45(5):055019.

[5] Deepanshi, Shivashankara K, Clifford G, Reyna M, Sameni R. ECG-Image-Kit: A Toolkit for Synthesis, Analysis, and Digitization of Electrocardiogram Images., 2024. URL `https://github.com/alphanumericslab/ecg-image-kit`.

[6] Reyna MA, Deepanshi, Weigle J, Koscova Z, Campbell K, Shivashankara KK, Saghafi S, Nikookar S, Motie-Shirazi M, Kiarashi Y, Seyedi S, Clifford GD, Sameni R. ECG-Image-Database: A dataset of ECG images with real-world imaging and scanning artifacts; a foundation for computerized ECG image digitization and analysis, 2024. URL `https://arxiv.org/abs/2409.16612`.

[7] Wu H, Patel KHK, Li X, Zhang B, Galazis C, Bajaj N, Sau A, Shi X, Sun L, Tao Y, et al. A fully-automated paper ECG digitisation algorithm using deep learning. Scientific Reports 2022;12(1):20963.

[8] Lampreave P, Jimenez-Perez G, Sanz I, Gomez A, Camara O. Towards assisted electrocardiogram interpretation using an ai-enabled augmented reality headset. Computer Methods in Biomechanics and Biomedical Engineering Imaging Visualization 2021;9(4):349–356.

[9] Isabel A, Jimenez-Perez G, Camara O, Silva E. Mobile app for the digitization and deep-learning-based classification of electrocardiogram printed records. In 2021 Computing in Cardiology (CinC), volume 48. IEEE, 2021; 1–4.

[10] Falk T, Mai D, Bensch R, Çiçek Ö, Abdulkadir A, Marrakchi Y, Böhm A, Deubner J, Jäckel Z, Seiwald K, et al. U-Net: deep learning for cell counting, detection, and morphometry. Nature methods 2019;16(1):67–70.

[11] Bradski G. The OpenCV Library. Dr Dobbs Journal of Software Tools 2000;.

Address for correspondence:

Álvaro José Bocanegra Pérez
Universitat Pompeu Fabra, Barcelona, Spain
alvaro.bocanegra@upf.edu