

Combining Hough Transform and Deep Learning Approaches to Reconstruct ECG Signals From Printouts

Felix Krones¹, Ben Walker², Terry Lyons², Adam Mahdi¹

¹ Oxford Internet Institute, University of Oxford, UK

² Mathematical Institute, University of Oxford, UK

Abstract

*This work presents our team’s (*SignalSavants*) winning contribution to the 2024 George B. Moody PhysioNet Challenge. The Challenge had two goals: reconstruct ECG signals from printouts and classify them for cardiac diseases. Our focus was the first task. Despite many ECGs being digitally recorded today, paper ECGs remain common throughout the world. Digitising them could help build more diverse datasets and enable automated analyses. However, the presence of varying recording standards and poor image quality requires a data-centric approach for developing robust models that can generalise effectively. Our approach combines the creation of a diverse training set, Hough transform to rotate images, a U-Net based segmentation model to identify individual signals, and mask vectorisation to reconstruct the signals. We assessed the performance of our models using the 10-fold stratified cross-validation (CV) split of 21,799 recordings proposed by the PTB-XL dataset. On the digitisation task, our model achieved an average CV signal-to-noise ratio of 17.02 and an official Challenge score of 12.15 on the hidden set, securing first place in the competition. Our study shows the challenges of building robust, generalisable, digitisation approaches. Such models require large amounts of resources (data, time, and computational power) but have great potential in diversifying the data available.*

1. Introduction

Cardiovascular diseases remain the leading cause of death globally [1], with electrocardiograms (ECGs) being essential for diagnosis and monitoring. Despite the rise of digital ECG devices, many ECGs are still recorded on paper, particularly in resource-limited settings [2]. Digitising these recordings could enhance datasets with greater demographic and temporal diversity, but the task is complicated by varying recording standards and image qualities.

Although digitisation efforts have advanced, they often

focus on a narrow set of images, leading to models that may not generalise well across different populations and recording conditions. The 2024 George B. Moody PhysioNet Challenge [2] addresses this gap by calling for robust algorithms to reconstruct ECG signals from printouts.

The 2024 Challenge included two tasks: recovering ECG signals from printouts and predicting cardiac diseases from these signals. Our team, *SignalSavants*, concentrated on the first task, with the goal of creating a versatile approach that can be trained and tested on diverse image sources by combining traditional image processing with deep learning techniques. We implemented an adaptable pipeline that includes image rotation using the Hough line transform, signal segmentation using deep learning, and mask vectorisation. We rigorously validated our models using cross-validation splits of 21,799 recordings from the PTB-XL dataset [3, 4], while creating multiple image versions for each signal to ensure robustness and reliability.

2. Methodology

2.1. Data

The PTB-XL dataset [3, 5] contains 21,799 12-lead ECG recordings from 18,869 patients. The data was collected from October 1989 to June 1996 and publicly released in 2019. For each of the recordings, we created at least four images using randomly applied augmentation options from the `ecg-image-kit` [6, 7]. We used a maximum rotation of 30 degrees and default settings for the other augmentation options. Figure 1 provides examples. Table 2 describes the randomly used augmentation options. We assumed the following standards across images, which we discuss further below: the same signal order, the same relative signal position, the same grid scale and the same amount of full-length signals.

2.2. Architecture

Our approach consisted of three major steps: image rotation, segmenting the image into individual signals, and

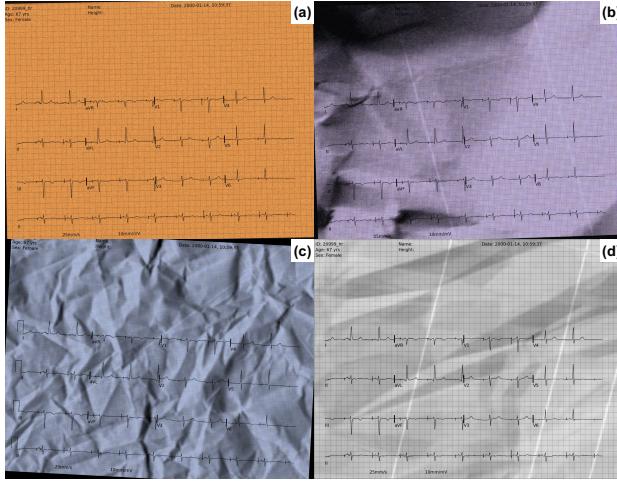


Figure 1. Four example images of signal 20999. (a) rotated by one degree, (b) rotated by two degrees with wrinkles and shadows, (c) rotated by three degrees with wrinkles and shadows, (d) not rotated with wrinkles and shadows.

vectorising the segmented pixels. Figure 2 is a schematic diagram outlining this approach.

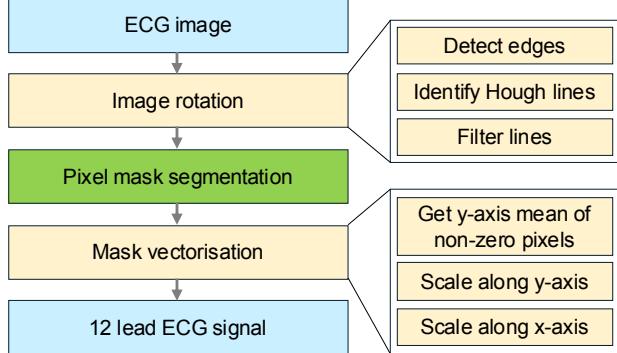


Figure 2. A schematic diagram of our model architecture. Blue: Data, Yellow: Rule-based engineering, Green: Deep learning.

Rotation. We rotated the ECG images to standardise the subsequent tasks. To achieve this, we used the `HoughLines()` method from the `opencv (cv2)` Python package. This method detects straight lines in an image by mapping edge points onto a polar coordinate system and finding intersections that indicate the presence of a line. We applied filters to these detected lines, considering only those within a specific angle range and ensuring a minimum number of parallel lines. Using the angles of these filtered lines, we calculated the rotation to align the ECG waveforms horizontally.

Segmentation. For the segmentation we used nnU-Net

[8], which was developed by the German Cancer Research Center. nnU-Net automatically configures the image processing and network architecture based on the data, using three types of parameters. First, it uses a set of fixed parameters (which remain constant due to their robust nature), such as the loss function, primary data augmentation strategies and learning rate. Second, rule-based parameters are adapted to the dataset using predefined heuristic rules, adjusting aspects like network topology and batch size. Finally, empirical parameters are determined through trial-and-error, refining elements like the best U-net configuration for the dataset and the postprocessing strategy.

Figure 3 shows a segmentation example. To create the masks we used two sets of pixels. The ones provided by the `ecg-image-kit` (called ‘sparse’) and an interpolated version that increased the pixel density (called ‘dense’).

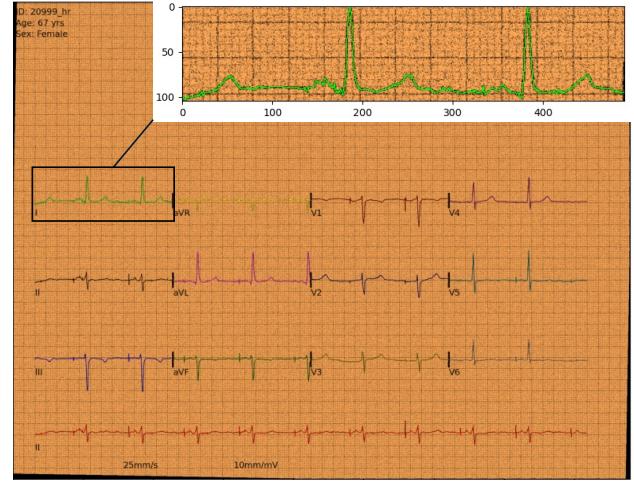


Figure 3. Example of predicted masks for signal 20999.

Vectorisation. In order to convert the mask, which is a matrix of pixel values, into a signal, the average y-position of the non-zero pixels in each column is used as signal value. Next, the signal is scaled to the correct amplitude using the grid size and resolution, which we assume to be constant. Finally, the starting time of the signal is determined based on the position of the signal on the ECG. Figure 4 shows an example of two leads from an ECG and the predicted signals from vectorising a mask of the signals.

2.3. Scoring

The metric used in the digitisation task was the signal-to-noise ratio (SNR) between the true signal y and the predicted signal \hat{y} , which can be calculated via

$$\text{SNR} = 10 \cdot \log_{10} \frac{\sum_i y_i^2}{\sum_i (\hat{y}_i - y_i)^2}. \quad (1)$$

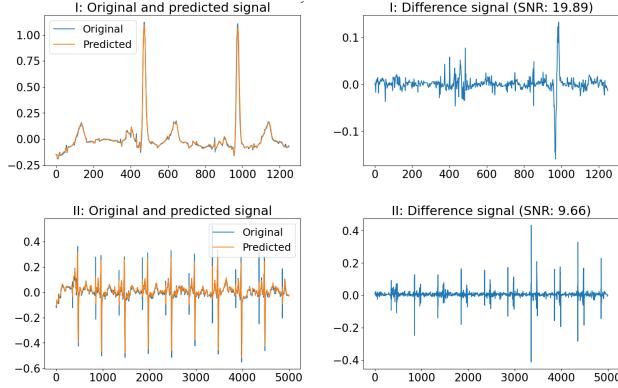


Figure 4. Vectorisation example for signal 20999. *Left:* original and predicted signal. *Right:* signal difference.

Letting the prediction be $\hat{y}_i = 0$ for all i provides a surprisingly strong baseline of $\text{SNR} = 0$ for this Challenge. This is due to the ECG signals having high and narrow peaks. During the Challenge, an update was released that aligns predicted and true signals up to a certain threshold along the x and y -axis [2].

To evaluate our models, we used the 10-fold splits suggested in [3] due to their stratified sampling approach.

3. Results

Table 1 shows the final challenge scores and Table 2 the results of different models for the digitisation task. Models M2 and M3, while in theory more advanced, were more prone to overfitting and needed more training data to achieve similar performance as M1. In the rotation step, we achieved a correct rotation prediction on over 99.7% of the images. Manually investigating a selection of the images where the rotation failed suggests that failure occurs when the colour schema and/or the augmentation are so extreme that determining the straight lines becomes very difficult. Vectorisation sometimes did not capture the peaks of the signals, as illustrated in Figure 4.

Task	Score	Rank
Digitization	SNR: 12.15	1/16
Classification	F-measure: -	-

Table 1. Signal-to-noise (SNR) ratio and F-measure of our team’s model on the hidden data for the digitization and classification (not participated) tasks, respectively.

In addition to the PTB-XL scans used in the official ranking, models were evaluated on a variety of additional image types [9]. Although we still achieved positive scores (0.51 to 4.93) on both colour and black-and-white scans of various paper qualities (deteriorated and clean), we scored

negatively (-1.76 to -0.72) on mobile phone photos and computer screenshots.

4. Discussion

In this study, we combined several common image processing approaches, such as the Hough transform for line detection, U-Nets for segmentation, and interpolation techniques for scaling. This method achieved the highest Challenge score, surpassing the next best-performing team by more than double.

However, the leaderboard only shows results for scans from the PTB-XL dataset, while the Challenge also evaluated entries on scans and mobile phone photos of images of different quality (e.g., clean, stained and deteriorated) as well as screenshots of computer monitors. No team achieved the highest score across all datasets. In fact, leaderboard performance dropped significantly on the other datasets (e.g., no team received a positive score for mobile phone photos of stained papers). Additionally, our local CV scores on various data augmentations were much higher than on the hidden sets. This performance gap highlights the challenge of achieving robust generalisation and the need for a training set that closely matches real-world conditions for reliable digitisation.

Some of our assumptions affected our results. Notably, all training examples had the signals in the same order and at the same position. While the segmentation model has the ability to learn the labels from the labels in the images, we assume that it simply assumed the same signal position for each prediction. We tested combining bounding boxes and optical character recognition (OCR) to be more independent of the signal position, but did not achieve robust results locally. Additionally, we assumed that each image contains only one full signal, which may not always be the case. For better generalisation, the segmentation model must be trained on examples with multiple full signals. We also assumed constant image resolution and grid scaling.

Our approach to ECG digitisation aligns with the multi-step approaches commonly reported in the literature [10–13]. Initially, we applied image pre-processing using the Hough transform [11, 12]. While many studies relied on pixel thresholds for segmentation [10–12], our objective was to improve generalisability by using a deep learning network (nnU-Net [8]), similar to the approach in [13], followed by vectorisation. However, unlike previous work [13], we applied our model to the entire image without prior pre-processing (except rotation), such as grid removal or detection of regions of interest. The key contributions of this approach are that a) it achieves much better results than other submissions to digitise ECG signals, and b) it can be easily applied to different data sources, whereas previous approaches are highly dependent on the specific style of the images used (e.g., assuming a partic-

Model	Train images	Augmentation	Mask	SNR official	SNR CV
M1	20k	Headers, calibration, wrinkles, rotation, cropping	Sparse	12.15	17.02
M2	45k	M1 + handwritten notes, QR-codes, various colours	Dense	9.42	12.58
M3	69k + 100 fine	Same as M2	Dense	4.20	15.27

Table 2. Digitisation models tested in this work. ‘SNR official’ is the leaderboard performance on the Challenge’s hidden set. ‘SNR CV’ is the average CV score. The column ‘Augmentation’ describes the randomly applied options from the `ecg-image-kit` during image generation. ‘Dense’ means an interpolated set of pixels was used for the masks (see 2.2). M3 was first trained locally on 69k images and then further fine-tuned on 100 of the hidden Challenge train data.

ular grid colour). As long as the training data mimic the deployment settings, our approach can be easily adjusted.

Developing methods to digitise images can enhance machine learning-supported screening systems by improving training through a wider range of data and helping to apply the models in more settings. However, making models robust and generalisable is not trivial. Achieving this requires a data-centric approach and vast resources.

Acknowledgements

We thank the Applied Computer Vision Lab of the Helmholtz Institute and the Medical Image Computing Devision at the DKFZ for making nnU-Net publicly available. FK was partially supported by the German tax payer through the Friedrich Naumann Foundation and BW and TL by the Hong Kong Innovation and Technology Commission and the Centre for Intelligent Multidimensional Data Analysis. TL was funded in part by the EPSRC [EP/S026347/1], The Alan Turing Institute [EP/N510129/1], the Defence and Security Programme, and the Office for National Statistics.

References

- [1] World Health Organisation. Cardiovascular diseases (CVDs). Geneva; 2021 Jun 11 [cited 2023 May 10]. Available from: [https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)).
- [2] Reyna MA, Deepanshi, Weigle J, Kosova Z, Elola A, Seyed S, Campbell K, Clifford GD, Sameni R. Digitization and Classification of ECG Images: The George B. Moody PhysioNet Challenge 2024 ;51:1–4.
- [3] Wagner P, Strodtboff N, Bousseljot R, Samek W, Schaeffter T. PTB-XL, a large publicly available electrocardiography dataset (version 1.0.3). <https://doi.org/10.13026/kfzx-aw45>, 2022.
- [4] Wagner P, Strodtboff N, Bousseljot RD, Kreiseler D, Lunze F, Samek W, Schaeffter T. PTB-XL: A Large Publicly Available ECG Dataset. Scientific Data 2020;
- [5] Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Compo-
- [6] 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;45:05019.
- [7] Deepanshi, Shivashankara K, Clifford GD, Reyna MA, Sameni R. ECG-Image-Kit: A Toolkit for Synthesis, Analysis, and Digitization of Electrocardiogram Images, January 2024.
- [8] Isensee F, Jaeger PF, Kohl SA, Petersen J, Maier-Hein KH. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nature Methods 2021;18(2):203–211.
- [9] Reyna MA, Deepanshi, Weigle J, Kosova Z, Campbell K, Shivashankara KK, Saghafi S, Nikookar S, Motie-Shirazi M, Kiarashi Y, Seyed 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>.
- [10] Fortune JD, Coppa NE, Haq KT, Patel H, Tereshchenko LG. Digitizing ECG image: new fully automated method and open-source software code. medRxiv 2021;2021–07.
- [11] Baydoun M, Safaty L, Abou Hassan OK, Ghaziri H, El Hajj A, Isma’el H. High precision digitization of paper-based ecg records: a step toward machine learning. IEEE Journal of Translational Engineering in Health and Medicine 2019;7:1–8.
- [12] 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.
- [13] 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 2020; 127:104077.

nents of a new research resource for complex physiologic signals. Circulation ;101(23):e215–e220.

- [6] 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;45:05019.
- [7] Deepanshi, Shivashankara K, Clifford GD, Reyna MA, Sameni R. ECG-Image-Kit: A Toolkit for Synthesis, Analysis, and Digitization of Electrocardiogram Images, January 2024.
- [8] Isensee F, Jaeger PF, Kohl SA, Petersen J, Maier-Hein KH. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nature Methods 2021;18(2):203–211.
- [9] Reyna MA, Deepanshi, Weigle J, Kosova Z, Campbell K, Shivashankara KK, Saghafi S, Nikookar S, Motie-Shirazi M, Kiarashi Y, Seyed 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>.
- [10] Fortune JD, Coppa NE, Haq KT, Patel H, Tereshchenko LG. Digitizing ECG image: new fully automated method and open-source software code. medRxiv 2021;2021–07.
- [11] Baydoun M, Safaty L, Abou Hassan OK, Ghaziri H, El Hajj A, Isma’el H. High precision digitization of paper-based ecg records: a step toward machine learning. IEEE Journal of Translational Engineering in Health and Medicine 2019;7:1–8.
- [12] 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.
- [13] 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 2020; 127:104077.

Address for correspondence:

Felix Krones
1 St Giles, Oxford, OX1 3JS, UK
felix.krones@oii.ox.ac.uk