Driving ECG Digitization - Using Techniques from Autonomous Driving to Detect Regions of Interest in ECG Images

Philip Hempel[1], Nicolai Spicher[1]

[1] Department of Medical Informatics, University Medical Center Göttingen, Germany

Abstract

Digitizing ECG images, i.e. reconstructing signal waveforms from paper printouts or screen captures could make AI-based diagnosis a reality, especially in the global south. We, team RoadRuNNers, transfer algorithms borrowed from autonomous driving — originally designed for road lane detection — to detect the region of interest (ROI) in ECG images and take them as input for a convolutional neural network (CNN). First, we invert the grayscale of the ECG image in order to make the ECG lines appear white before a dark background, similar to a road lane. Subsequently, we adopt open source scripts that iteratively detect white lines and their orientation by randomly choosing thresholds for binarization, Gaussian blur kernel size, and Canny edge detection. The detected $4-6$ *ROIs are assumed to show ECG signals when they are in parallel and equidistant to each other and span at least half of the image size. The ROIs then serve as input for a CNN that processes the* 4 *binary images to reconstruct* $4 \times 10s$ *of ECG signal. Our approach in 'Digitization and Classification of ECG Images: The George B. Moody PhysioNet Challenge 2024' achieved a digitization score of* -0.071 *(rank 10/16) on the hidden test data. Despite improvable performance levels, we demonstrated the potential of adapting techniques from autonomous driving for ROI detection.*

1. Introduction

Electrocardiograms (ECGs) are a basic and non-invasive examination technique in medical care. Despite advances in digitization, many ECGs are still printed on paper or displayed as pixel-based images on screens, especially in resource-limited settings. Hence, digitizing ECG images holds the potential to democratize access to AI-driven diagnostics globally, particularly in the Global South where healthcare resources are often scarce.

The objective of the 2024 George B. Moody challenge [1–3] is to reconstruct the signal from ECG images with varying quality and format. Traditional methods of signal reconstruction have been reliant on manual or semi-automated processes[4], which are time-consuming and prone to errors. Our team, RoadRuNNers, propose a novel approach that leverages techniques originally developed for autonomous driving, particularly road lane detection, to fully automatically identify regions of interest (ROIs).

Road lane detection is a crucial component of autonomous driving, designed to recognize and track road lanes in real-time using image processing algorithms. These algorithms identify straight lines in images based on specific characteristics, such as color contrast and alignment [5, 6]. Detection needs to be efficient as it is carried out on large number of images which are often noisy due to the dynamic environment [7–9].

In the context of ECG digitization, we hypothesized that these techniques could be repurposed to detect ROIs containing ECG waveforms, which – like road lanes – are linear features with consistent orientation and spacing within an image.

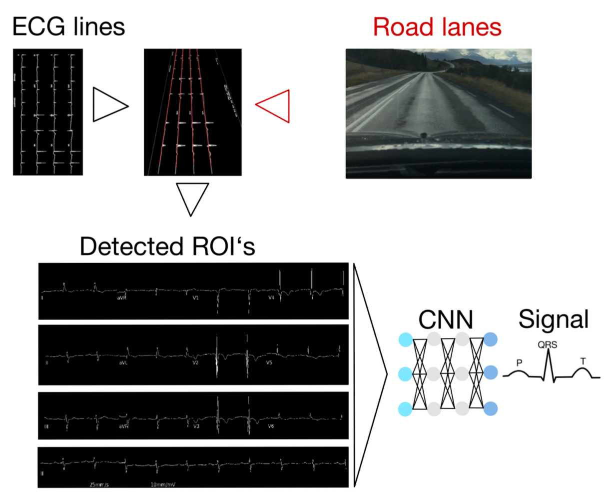

Figure 1. Graphical abstract: Road lane detection and ECG digitization face similar challenges. We hypothesize that algorithms established for road lane detection can be adapted to detect ROIs containing the ECG waveforms.

 ISSN: 2325-887X DOI: 10.22489/CinC.2024.247

2. Methods

We first create a training dataset according to the challenge instruction [1, 3] using the PTB-XL dataset [10] and the provided toolkit [3, 11] with randomly chosen parameters. Subsequently, by using `OpenCV2` [12], parallelism and equidistant spacing of the ECG baselines are employed to perform the iterative detection of ROIs, which encompasses a series of steps:

2.1. ECG Image Preprocessing

To facilitate the detection of ECG waveforms, each ECG image is initially converted to grayscale and inverted. This inversion results in the ECG lines appearing white against a dark background, which closely resembles the appearance of road lanes in typical lane detection tasks [13, 14].

2.2. Binarization and Edge Detection

Following the inversion, the image undergoes binarization, where a threshold is iteratively adjusted to create a binary image that distinctly separates the ECG lines from the background. A Gaussian blur is applied to the binary image to smooth out noise and enhance the continuity of the lines [15]. Subsequently, we perform edge detection [15] using the Canny edge detector, with thresholds optimized to capture the linear features of the ECG waveforms. Parameters were manually defined by heuristic optimization.

2.3. Contour Detection, Orientation Analysis, and Slope Correction

We identify contours within the edge-detected image, that represent potential ECG lanes. Each contour's orientation and length are calculated using the Hough Line Transform, a method commonly used in road lane detection [16]. Contours that exhibit a straight orientation and a consistent slope are retained, while those deviating from the median orientation are discarded. This filtering step ensures that the detected contours are parallel and align with the expected horizontal orientation of the ECG baseline.

To correct for any tilt in the ECG lines, we apply slope correction. If the detected median slope of the contours deviates by more than $\pm 5^\circ$ from the horizontal, the entire image is rotated by the corresponding angle to align the contours horizontally. This correction step ensures that the ECG waveforms are properly aligned, facilitating accurate subsequent analysis. Following slope correction, we reevalute the contours to confirm they meet the horizontal alignment criteria, to further refine the detected ECG lanes.

2.3.1. Grouping and Equidistance Analysis

We group the selected contours based on their spatial alignment and proximity. The y-coordinates of the contours' centers are used to form subgroups that represent potential ECG lanes. A tolerance level is set to accommodate minor variations in alignment. We calculate the equidistance between these groups, with only those maintaining consistent spacing and spanning at least half the width of the image being considered valid ROIs.

2.4. ROI Extraction

For each image, the ROIs are extracted by selecting the areas surrounding the detected ECG lanes, to ensure that these regions encompass the relevant ECG signals. We resize the extracted ROIs to a fixed dimension of 2000×300 pixels, and all four are used as input for the CNN.

The parameters, including the paths to the original ECG records, image paths, detected ECG lanes, lane distances, binary thresholds, and labels, are saved in a serialized format using `pickle`. This allows for efficient reuse during training and testing without reprocessing the raw data in every epoch, to reduce computation time. When constructing datasets for model training, we load these saved parameters and use them to generate consistent input tensors for the CNN, ensuring that the model receives data in a standardized format across different training iterations.

2.5. Signal Reconstruction Using CNN

The extracted four ROIs from the ECG images are fed into a custom CNN to reconstruct the ECG signals. We apply a custom loss function during training to optimize the model, which aims to improve the reconstruction of the peaks of the ECG signals:

$$\text{WeightedLoss} = \frac{1}{N}\sum_{i=1}^{N} L_{\text{smooth\_L1}}(\text{outputs}_i, \text{targets}_i) \cdot w_i$$

$$w_i = \begin{cases} 5, & |\text{targets}_i| > 0.2 \\ 1, & \text{otherwise} \end{cases}$$

where:

- WeightedLoss is the weighted average loss over all elements in the batch.
- N is the total number of elements in the batch.
- $L_{\text{smooth\_L1}}(\text{outputs}_i, \text{targets}_i)$ is the Smooth L1 Loss between the i-th output and target.
- w_i is the weight applied to the i-th element, where:
  - $w_i = 5$ if $|\text{targets}_i| > 0.2$ (for peaks),
  - $w_i = 1$ otherwise.

3. Results

Fig. 2 depicts the outcomes of our ROI detection pipeline. The combination of adaptive thresholds with the assessment of ECG-specific orientation and placement results in robust ROI detection (Fig. 2D).

However, the four detected ROI that serve as input for the trained CNN did only reconstruct a general ECG baseline which result in the challenge score for the official phase depicted in Tab. 1.

Table 1. Signal-to-noise ratio (SNR) achieved by our team 'RoadRuNNers' on the hidden test data for the **digitization task**.

Task	**Score**	**Rank**
Digitization	-0.071	10 / 16
Classification	-	-

4. Discussion and Conclusion

Our results demonstrate the efficacy of the ROI detection approach, which effectively isolated relevant ECG segments. However, the limitations in the CNN's ability to reconstruct the ECG signals – visible in Tab.1 –, suggest that the complexity of reconstructing ECG waveforms from these segments may require a more sophisticated or alternative modeling approach. Despite extensive efforts to adapt the CNN, including the use of custom loss functions, attention mechanisms, and various architectural adjustments, the results revealed that these modifications were insufficient to overcome the inherent difficulties in the task.

To address the challenges faced in CNN-based ECG reconstruction, future work could consider the following:

- **Processing ROIs Individually:** Rather than aggregating ROIs, processing each segment individually might allow for more detailed and accurate reconstruction of ECG waveforms, reducing the complexity of the task for the CNN and potentially improving its performance.
- **Replacing CNN Architecture:** Changing the CNN architecture to better capture the specific ECG characteristics. For instance, incorporating domain-specific knowledge into the model design or experimenting with different types of neural networks, such as recurrent networks or transformers, could enhance signal reconstruction.

While road lane detection and ECG signal digitization appear to be disparate fields, they in fact share a number of significant challenges. This common ground makes techniques developed in one area potentially applicable to the other. This involves the analysis of sequential patterns in visual data [13], identification of noise, and the assessment of distortion [14]. In the context of road lanes, noise may result from factors such as shadows, occlusions, or lighting

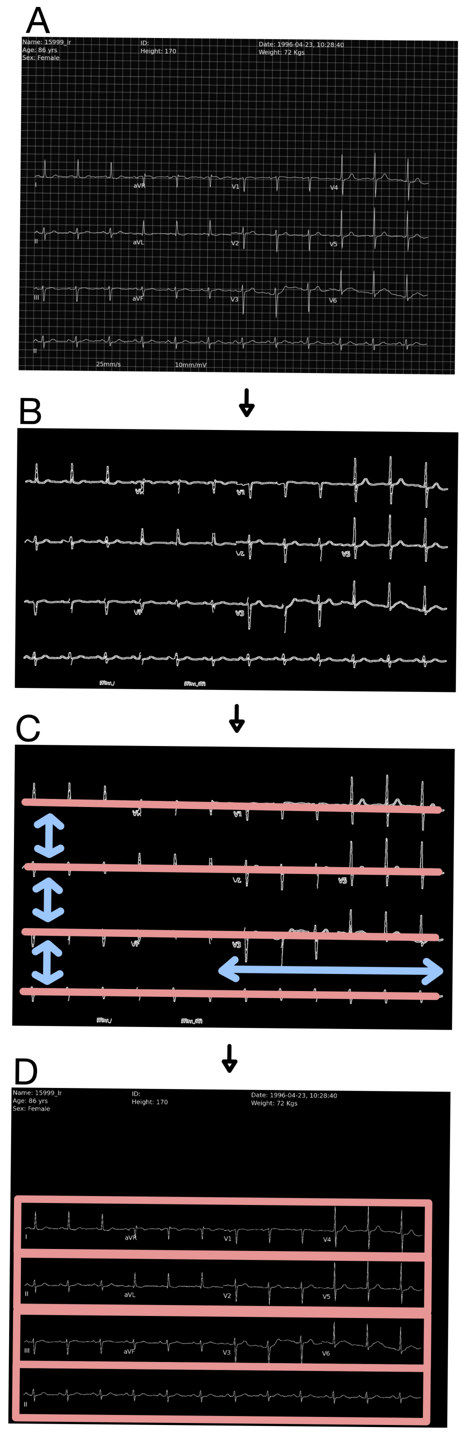

Figure 2. Steps of ROI detection. **A:** Preprocessed image with inverted colors **B:** Iteratively chosen threshold for binarization and edge detection **C:** Found structures are checked to be parallel, equidistant and span at least the half of the image **D:** Detected ROIs which are fed to the CNN.

variations. In ECGs, noise may originate from artifacts, baseline wander, or other distortions in the image.

Given the high stakes involved in both autonomous driving and medical applications like ECG digitization, the intense and continuous advancements in developing robust algorithms for self-driving systems could be highly valuable. The sophisticated solutions designed to ensure safety and accuracy in autonomous vehicles may be effectively adapted and repurposed to meet the critical demands of ECG digitization. By leveraging the technologies and methodologies that drive reliability in autonomous systems, we can potentially enhance the precision and performance of medical digitization tasks, where the margin for error is equally narrow.

While the ROI detection achieved notable success, the challenges encountered in signal reconstruction highlight the need for further refinement of the CNN or the exploration of alternative approaches. Nevertheless, the parallels between road lane detection and ECG digitization suggest that adapting techniques from autonomous driving could offer valuable insights and improvements, particularly in addressing the complexity and noise inherent in ECG images.

Code Availability

Code contributed to the challenge is openly available at: `https://gitlab.gwdg.de/medinfpub/biosignal-processing-group/georgebmoodychallenge2024_roadrunners.git`

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] Goldberger AL, Amaral LAN, 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.

[3] 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`.

[4] Fortune JD, Coppa NE, Haq KT, Patel H, Tereshchenko LG. Digitizing ECG image: A new method and open-source software code. Computer Methods and Programs in Biomedicine 2022;221:106890.

[5] Zhou Cc. Ultra-fast lane detection v2: Open-source code for high-speed lane detection. `https://github.com/cfzd/Ultra-Fast-Lane-Detection-v2/tree/master`, 2022. Accessed: August 28, 2024.

[6] Zou Q. Robust lane detection: Code repository for robust lane detection research. `https://github.com/qinnzou/Robust-Lane-Detection`, 2022. Accessed: August 28, 2024.

[7] Taherian S. Lane detection. `https://github.com/shayantaherian/Lane-Detection`, 2020. Accessed: August 28, 2024.

[8] Razack S. Vehicle Lane Detection using YOLOv2: Open-source implementation for vehicle lane detection. `https://github.com/shazraz/Vehicle-Lane-Detection-YOLOv2`, 2017. Accessed: August 28, 2024.

[9] Team TD. LaneDet: An Open-source Framework for Lane Detection in Autonomous Driving. `https://github.com/Turoad/lanedet`, 2021. Accessed: August 28, 2024.

[10] Wagner P, Strodthoff N, Bousseljot RD, Kreiseler D, Lunze K, Samek W, Schaeffter T. PTB-XL, a large publicly available electrocardiography dataset. Scientific Data 2020; 7(1):1–15.

[11] 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.

[12] Bradski G. The OpenCV Library: An Open-source Computer Vision Framework. `https://www.drdobbs.com/open-source/the-opencv-library/184404319`, 2000. Accessed: August 28, 2024.

[13] Yim YU, Oh SY. Three-feature based automatic lane detection algorithm (tfalda) for autonomous driving. IEEE Transactions on Intelligent Transportation Systems 2003; 4(4):219–225.

[14] Kemsaram N, Das A, Dubbelman G. An integrated framework for autonomous driving: Object detection, lane detection, and free space detection. In 2019 Third World Conference on Smart Trends in Systems Security and Sustainability (WorldS4). IEEE, 2019; 260–265.

[15] Devane V, Sahane G, Khairmode H, Datkhile G. Lane detection techniques using image processing. In ITM Web of Conferences, volume 40. EDP Sciences, 2021; 03011.

[16] Aminuddin NS, Ibrahim MM, Ali NM, Radzi SA, Saad WHM, Darsono AM. A new approach to highway lane detection by using hough transform technique. Journal of Information and Communication Technology 2017; 16(2):244–260.

Address for correspondence:

Philip Hempel
Robert-Koch-Str. 40, 37075 Goettingen, Germany
philip.hempel@med.uni-goettingen.de