End-to-End ECG Digitisation with Soft Segmentation and BiLSTM Signal Reconstruction

Adrian McIntosh[1], Mojtaba Lashgari[1], Longa Maimbo[2], Abhirup Banerjee[1]
[1] Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford
[2] Department of Electrical Engineering, University of Cape Town

Introduction: ECGs generate vast amounts of data annually with strong potential for AI-driven population studies, yet, much of this data remains in paper format, limiting computational use. Digitising these records, especially via images, is challenging due to uneven illumination, noise, and artefacts, such as handwritten annotations that can obscure the signal.

Aim: To develop a novel end-to-end paper ECG digitisation pipeline that is robust to realistic input distortions and artefacts.

Methods: An end-to-end pipeline is proposed following three stages: image rectification, probability heatmap segmentation, and a Bidirectional Long Short-Term Memory (BiLSTM)-based signal reader. Unlike conventional approaches that utilise rule-based trace-following methods, the proposed method converts segmentation outputs using sequential modelling, enabling robust reconstruction in the presence of realistic input distortions and artefacts. The pipeline is demonstrated on the PhysioNet 2025 ECG Digitisation Challenge dataset, comprising 997 12-lead ECG signals, each rendered as nine corresponding synthetic images with varied capture conditions.

Results: Evaluated end-to-end on a held-out test set of 1,323 images (generated from 147 ground-truth digital ECG signals), the pipeline achieved a mean signal-to-noise ratio of 16.89 $\pm$ 3.11 dB with consistent performance across image distortion types, demonstrating the potential of this fully automated ECG digitisation framework.

Conclusion: This work demonstrates that a fully automated, learning-based pipeline can robustly digitise paper ECGs under realistic and varied capture conditions, unlocking previously inaccessible records and automated analysis.

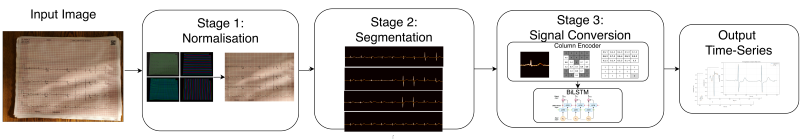

Figure 1. An overview of the ECG digitisation pipeline.