Advancing Cardiac Care through Digitization and Classification of Electrocardiograms

Bjørn-Jostein Singstad, Jesper Ravn, Henrik Schirmer, Arian Ranjbar

Medical Technology and E-health, Akershus University Hospital
Lørenskog, Norway

Introduction: The ECG is a non-invasive method for measuring the heart's electrical activity, helping doctors to diagnose and prognosticate cardiovascular diseases. Since the string galvanometer in 1895, this technique has evolved significantly to include digital waveforms, automatic interpretation, and integration into wearables. Despite these advancements, the majority of ECGs are still printed on paper rather than stored directly as digital waveforms. This limits their use with modern AI development which has shown rapid progress recently in interpreting ECGs and diagnosing various cardiovascular conditions. Digitizing paper-ECGs could enable combining them with digitally stored ones to enhance AI algorithms further.

Method: Waveform data from the PTB-XL database was used to generate synthetic paper-ECGs with an open-source Python software package called ECG-Image-Kit. Furthermore, a paper-ECG digitization pipeline was created consisting of lead detection, thresholding, binarization, 50Hz noise filtering, baseline wander removal,and resampling followed by voltage scaling and zero-padding. Subsequently, a 1-dimensional convolutional neural network was trained on the digitized data to classify the ECG as normal or abnormal. The model was trained on the digitized 12-lead ECGs upsampled to 100Hz and zero-padded to 10 seconds. The model ran for 12 epochs with a batch size $= 30$ and categorical cross-entropy as loss function.

Result: Our team; Ahus AI Lab, developed a digitization and classification model and the digitization achieved a signal-to-noise ratio of -0.28 while the classification model achieved an F1-score of 0.860 ± 0.007 (10-fold cross-validation). However, the code failed to run on the organizers' hidden validation set, resulting in no scores being achieved.

Conclusion: Despite our promising results we should also bear in mind that the development and validation of the digitalization model was done on synthetically generated paper-ECGs. Real scanned paper-ECGs may show larger variations due to artifacts. Future advancements should focus on evaluating performance with real paper-ECGs.