

Investigating the Robustness of Deep Learning to Electrocardiogram Noise

Jenny Venton, Philip J Aston

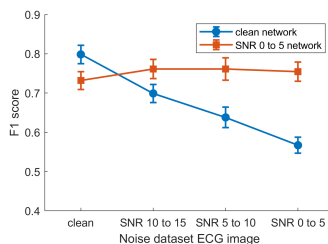
Data Science, National Physical Laboratory
Teddington, UK

Aim: This study examined how different signal-to-noise ratio (SNR) levels of physiological electrocardiogram (ECG) noise in deep learning training data influenced network robustness to physiological noise. Physiological noise on an ECG is a known source of error in ECG diagnosis and interpretation and has been noted to affect automated (machine learning) detection of ECG abnormalities.

Method: An ECG dataset containing 2678 signals was filtered to remove as much noise as possible. Prerecorded physiological ECG noise was then added to copies of the clean dataset, producing four datasets with four different noise levels. In order of cleanest to noisiest these were i) clean, ii) SNR between 10 dB and 15 dB, iii) SNR between 5 dB and 10 dB, iv) SNR between 0 dB and 5 dB. Scalogram images were generated for all signals in each dataset using a continuous wavelet transform. A pretrained ResNet-50 model was adapted to classify the ECG images for each dataset using transfer learning. Finally, each trained network was used to classify unseen ECG images from all noise level test datasets, to understand how including noise in the training data affects network robustness to noise in the input data.

Results: The network trained on clean ECG images performed best when classifying clean images (F1 score 0.80), but performance decreased dramatically for classifying the noisiest images (F1 score 0.57, see Figure). However, although the network trained on the noisiest images did not perform as well on the clean images (F1 score 0.73), its performance across the noise images improved and remained fairly consistent (F1 score ~ 0.76 , see Figure).

Conclusion: The inclusion of physiological ECG noise in deep network training data improves the robustness of the network to different levels of noise in the input data.



Network performance when classifying the four datasets. All SNR units are dB.