Matrix Multiplication and Baseline Analysis Methods to Classify ECGs in the 2011 Physionet/CinC Challenge

Arie C Maan*, Erik W van Zwet, Suzanne M M Oliveira-Martens, Sumche Man, Martin J Schalij, Ernst E van der Wall and Cees A Swenne

Leiden, Netherlands

We present an algorithm for classifying ECGs, recorded in remote areas by inexperienced persons, that may or may not be suited for expert reviewing after trans-telefonic transmission of the signal. The work presented here was initiated by the 2011 Physionet/Computing in Cardiology Challenge in order to improve the quality of ECGs recorded in rural areas with limited access to well equipped hospitals, but, obviously, the algorithms should also be applicable for automated in-hospital quality control of the many clinically recorded ECGs. ECGs should fulfill two criteria: 1) any noise (possible causes: baseline drift, high-frequency noise, low signal-to-noise ratio) present in the signal should not interfere with important features like morphology of P-QRS-T and 2) the electrodes should be placed correctly. The first criterium was addressed by 1A) determining the residue after high-pass filtering to assess the baseline stability, 1B) determining the amount of high-frequency noise and 1C) by cross correlation of the signal. The second criterium was addressed by 2A) mathematically transforming the ECG into a vector cardiogram (VCG) using the Kors matrix followed by a back transformation (reconstruction) into a 12 lead ECG by using an optimized inverse matrix and determining, for each lead, the correlation between the original and the reconstructed ECG and 2B) calculating the Singular Value Decomposition (SVD) of the signal. A learning set of 1000 ECGs was provided, of which 606 were acceptable according to a panel of 3 experts, and 394 were not. Using these data we found a correct prediction of 86.3%, 74.1% and 77.6% for 1A, 1B and 1C respectively, and of 85.4% and 75.8% for 2A and 2B, respectively. Preliminary results from combinations of these methods yielded an 89.1% correct prediction for the learning set and a correct prediction of 85.0% for the test set.