Signal Quality Indices and Data Fusion for Determining Acceptability of Electrocardiograms Collected in Noisy Ambulatory Environments

Qiao Li, D Lopez and Gari Clifford*

Oxford, Oxfordshire, United Kingdom

An algorithm is described to detect poor quality ECGs collected in low-resource environments as part of the PhysioNet/Computing in Cardiology Challenge 2011 'Improving the quality of ECGs collected using mobile phones'. The algorithm is based on previously published signal quality metrics designed for intensive care monitoring. The algorithms have been adapted for use on short (10s) 12-lead ECGs. The metrics quantify spectral energy distribution, higher order moments and inter-channel and inter-algorithm agreement. Five metrics are produced for each channel (60 features in all) and presented to a Support Vector Machine (SVM) for training on the provided labeled data (Set-a) for the challenge. (Binary labels were available, indicating whether the data were acceptable or unacceptable for clinical interpretation.) The 1000 subjects in Set-a data were split into two balanced sets (training data and test data) with 480 subjects in the training set. The SVM provided a classification accuracy of 87.71% on the training data and 85.0% on the test data. The trained SVM was then applied to an unlabeled data set (Set-b), and the binary classification results (acceptable or unacceptable) were submitted as an entry into the PhysioNet Computing in Cardiology Competition 2011. Errors in the classification algorithm were due to ambiguous labels in the training data, single lead inversions carrying little weight, and the classification of arrhythmias as unacceptable. These issues will be addressed in the coming months. Future improvements of the algorithm are also described.