Rule-Based methods and Deep Learning Networks for Automatic Classification of ECG

Giovanni Bortolan, Ivaylo Christov

Institute of Neurosciences IN-CNR, Padova, Italy

Objective: The objective of the material is to explore the potentiality of combining a classical rule-based method with a Deep Learning method for automatic classification of ECG in the framework of the PhysioNet Challenge 2020.

Materials: The training set is a public database collected in China and consists of 6877 (3699 males and 3178 females) 12-leads ECG recordings lasting from 6 seconds to 60 seconds. Each ECG recording has one or more diagnostic labels: Normal sinus rhythm, Atrial Fibrillation (AF), First-degree Atrio Ventricular Block (I-AVB), Premature Atrial and Ventricular Complex (PAC, PVC), Left and Right Bundle Branch Block (LBBB, RBBB), ST-segment depression and elevation (STD, STE).

Method: The rule-based method is using morphological and time-frequency ECG descriptors, characterizing each diagnostic labels. These rules have been extracted from the knowledge-base of a physician, with no direct learning procedure in this phase. A refinement in the training database will be tested in the second phase. For example, more specific limits for PVC, PAC, STE or STD can be usefully extracted from the training set.

The Deep Learning method consider both raw ECG signals and median beat signals. These data are processed by continuous wavelet transform analysis obtaining a time-frequency domain representation, with the generation of specific images. These images are then used for training Convolutional Neural Networks for ECG diagnostic classification. Two pre-trained CNN for image classification have been initially used, GoogleNet and SqueezeNet networks, and the Alex Network will be tested in the second phase.

Results: Unofficial results of the classification accuracy on an independent ECG testing set, are reported in the Table. These results suggest the possibility to increase the accuracy with a more robust learning process for DL and the use of appropriate rule-refinement for benefitting of the training set. In addition, the possibility to explore the potentiality of combining the two methods will be quantified and studied in the second phase. The poor accuracy of the two DL methods only in the testing process (e.g. F_2: 0.070 vs 0.575 with cross-validation) suggests the presence of a systematic deviation/error in the processing phase, which will be deeply examined in the next phase.

Results of the Challenge

Method	F_2	G_2
Rule-based	0.530	0.283
DL GoogleNet	0.070	0.028
DL Squeeze N	0.039	0.011