

1 Abstract

One the widely used heart-condition examining method is phonocardiogram (PCG), a plot of high-fidelity recording of the sounds and murmurs made by the heart with the help of the machine called the phonocardiograph. By analyzing PCG record doctors can detect abnormal heart activity that classifies for special treatment. Such classification requires expert knowledge and can be very time consuming. By introducing programs for initial PCG classification we can reduce time spent on PCG analyzing.

Classification of heart activity in PCG recordings are usually solved in three steps: segmentation, feature extraction and classification. We base our work on the prepared sample MATLAB entry for Physionet competition, in which state-of-the-art segmentation and simple feature extraction have already been prepared. We present an enhanced feature extraction and classification. Effectiveness of some of the best known classification methods is compared by taking their accuracy and speed into account. This includes SVM, Random Forests, KNN and HMM methods. We use grid search method to find and set best parameters for each classifier. All sources are written in MATLAB.

Samples used in this work are provided by Physionet challenge creators and are accessible during its duration. Whole records set consisted of 3126 files containing PCG recordings grouped into 6 folders. The validation package is used as a training set during classifiers' learning process. Audio files available within additional .zip file named *training* are divided into 5 folders named: *training - a*, *training - b*, ..., *training - e* and are used as test sets for which classifier results have been compared. All of mentioned audio examples have already been classified by an expert and the results for their classification is known.