

PCG Classification using a Neural Network Approach

Iga Grzegorzczuk, Mateusz Soliński, Michał Łeppek, Anna Perka, Jacek Rosiński, Joanna Rymko, Katarzyna Stępień, Jan Gieraltowski

Faculty of Physics, Warsaw University of Technology, Warsaw, Poland

Abstract

Phonocardiography (PCG) is the one of non-invasive ways to diagnose condition of human heart. The mechanics of heart muscle contractions and closure of the heart valves generates vibrations audible as sounds and murmurs, which can be analysed by qualified cardiologists. Developing an accurate algorithm to determine whether patients' heart works properly or should be referred to an expert for further diagnosis would significantly improve the quality of healthcare system. It would allow to perform less unnecessary, expensive and time consuming examinations.

The analysed data consisted of PCG recordings from the training set provided by the organizers of the PhysioNet Challenge 2016. Its length varied from several to 120 seconds.

We propose the machine learning algorithm based on neural networks. The segmentation of the PCG signals is performed with algorithm based on Hidden Markov Model. Whereas, the features necessary to define whether the signal looks normal or should be further analysed were carefully chosen by our team and belonged to time domain, ordinate axis or frequency domain group. The great emphasis was put on the statistical features representing the characteristics of the signal. Their optimal values were found during the process of learning of our algorithm.

The best overall score we achieved in the official phase of the PhysioNet Challenge 2016 is 0.79 with specificity 0.76 and sensitivity 0.81.

1. Introduction

Auscultation is one of the examination techniques most commonly used by the doctors during both hospital and house visits. It allows to diagnose a wide range of pathologic cardiac conditions from arrhythmias to valve diseases or heart failure. The heart sounds audible in PCG recordings result from the mechanical activity of the heart muscle. Contractions of ventricles and atriums cause the closure of the respective valves. They start to

vibrate with their natural frequency, what can be heard via stethoscope. Apart from the correct heart sounds such as S1 and S2 the pathological murmurs can be heard. Depending on their localization different pathologies and structural defects can be diagnosed.

We analysed over 3 000 heart sound recordings from the PhysioNet Challenge 2016 training set. It contained very short PCG signals lasting between 5 and 120 seconds. Each signal was already labelled as normal or requiring the expert consult. Our task was to create an algorithm which will divide set of signals into such two groups. All signals were provided as *.wav format and were sampled at rate of 2 000 Hz. Further information about the database and challenge rules can be found in [1].

The core of our algorithm is the neural network with 48 features. To assure that our network will learn only on good quality signals (because the bad quality ones would distort the values of features on which the network was trained), we excluded from our training set presented to the network all of the signals that organizers marked as uncertain.

Based on the set of the features on which the network was trained it decides whether the analyzed signal should be classified as normal or abnormal, which was the main task of the PhysioNet/Computing in Cardiology Challenge 2016.

2. Methods and results

2.1. Preparation and analysis of the signals

One of the base steps performed by our algorithm is the segmentation of PCG signal, which is necessary to calculate the chosen features. We used the state-of-art algorithm utilizing Hidden Markov Model [2], which was further developed by Springer [3]. It allowed us to determine all of the features describing the consecutive phases (S1, S2, Systole, Diastole) in time and ordinates axis.

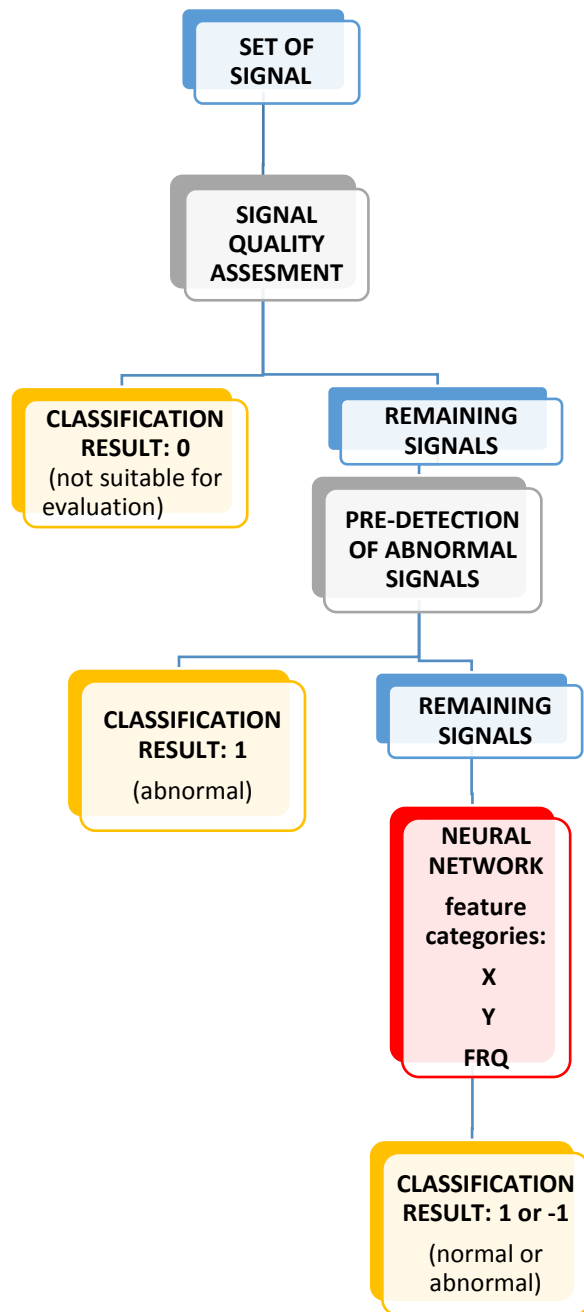


Fig.1 Flow chart of PCG classification algorithm.

2.2. Pre-neural network classification

As shown on the Figure 1, before the neural network was applied we excluded some recordings from further analysis by already classifying them with set of criteria to assess the signal quality and by pre-classification of abnormal signals, which are described in detail below.

Assessment of signal quality

We assessed the quality of the PCG signal as a first, starting step. At the beginning we used the method for preliminary detection of peaks in signal, which are inclined to occur during the S1 or S2 phase (whichever have higher amplitude in PCG cycle). We calculated the energy of the normalized signal (the normalization was performed separately in each of 25 windows of the signal) as described in [4]. Next we determined the peaks with Gierałtowski et al. method [5] based on the slope detection in the signal. In the next part of the method, we calculate Wavelet coefficients of the whole signal by using Daubechies-2 wavelet at second decomposition level (according to [6]) and use these coefficients as an input signal to later evaluations. We created three sets of criteria for assessing signal quality. If the signal did not meet at least one criterion it was classified as uncertain. First criterion: RMSSD (root mean square of successive differences) of the signal must be lower or equal to 0.026 [arbitrary units]. Moreover, the number of times which the signal intersected the horizontal line determined by 0.85 quantile of values of that signal divided by the signal length must be lower than 0.06. Second criterion: we analyzed the signal in 2200 ms length moving windows with overlap 25% and checked how many peaks are detected in each window. We assigned 1 to each window containing 2-4 peaks and 0 in other cases. The criterion was met when the number of scores equal to 1 is 65% of the all notes calculated in each window. Third criterion was similar to the first one, but in this case there was calculated the value of 0.58 percentile of signal and the number of detected peaks was extracted from the whole number of intersections. According to these rules, the number of intersections must be fewer than 18. Note that the second and the third criterion were checked depending on the number of detected peaks per one second of the signal. If that value was higher than 1.1 s, signal had to meet only the second criterion. Otherwise, only the third criterion was checked.

Pre-classification of abnormal signals

The second step of the algorithm classified the recordings as abnormal based on their systole and diastole phases without using the neural networks. This approach was chosen to eliminate evidently abnormal recordings, which might be classified without complex methods. This way we could save time and assure that by limiting the variety of the signals, they will not distort values of the features presented to the neural network. We created the method providing high precision of detection abnormal cases (91%) but with relatively low specificity (11% of the whole set of abnormal signals). As the input to this method we used the PCG signals and the assigned states obtained from segmentation algorithm. We calculated three features, separately for systole and diastole phase and for both phases joined in one signal.

We normalized each beat of the signal by mean and by the difference between maximum and minimum value. Then, in each of the segmented PCG cycles: RMSSD, SD1 of the Poincare plot and the number of zero crossings of the signal (normalized by the signal length) were calculated. Next, we determined the difference between these features calculated for systole and diastole phases divided by the values obtained from joined phases. Finally, we analysed the values of the features in each beat and assign 1 for beats which are exceeding at least one of established thresholds (0.8 for RMSSD, 0.6 for zero crossing and 0.8 for SD1) and 0 otherwise. If the number of 1 notes divided by the number of whole detected beats is higher than the second set of thresholds (0.8 for RMSSD, 0.8 for zero crossing and 0.7 for SD1) the signal is classified as abnormal. The values of the thresholds were established by using exhaustive search method.

2.3 Features to neural network

After performing pre-neural classification, neural network is responsible for recognizing normal and abnormal signals. The outcome depended on the features extracted from PCG signals. Three categories of features can be distinguished, as presented in Fig.1:

X – features based on parameters obtained with signal observation by time axis,

Y – features based on parameters obtained with signal observation by ordinates axis,

FRQ – features based on calculations on signal parameters obtained by examination in frequency domain.

Description of the parameters included in each of the categories is presented below:

X:

In the group of time domain features we concentrated on those carrying information about regularity of length of the consecutive phases (S1, Systole, S2 and Diastole). We described these parameters by the mean and standard deviation calculated in the whole signal. The abnormalities in proposed features might indicate additional sounds, which in adults signify pathologies and should be classified as abnormal. Additionally, we enclosed feature determining the ratio of length of S1 and S2 interval. Moreover, we assumed that the summarised length of S1, Systole and S2 phases can represent the QT interval in each heart beat and we calculated its standard deviation among the signal. Finally, we included some of the parameters proposed by the organizers in Sample Entry which regarded lengths of consecutive phases.

Y:

To the second group were assigned features based on parameters from ordinates axis obtained mostly by the

close observation of the signal. In this part we had taken into account the characteristics, which described mean value of the mean absolute amplitude ratios between systole/diastole period and S1/S2 period in each cycle. We also included the standard deviation value of the mean absolute amplitude ratios between systole/diastole period and S1/S2 period in each heartbeat.

FRQ:

The last of the proposed groups included parameters calculated in frequency domain. The first pair of parameters represented the power of S1 and S2 intervals respectively. They were calculated as the sum of the absolute values of signal in this phases. Then we determined the root square mean of the squares of the successive differences between adjacent intervals (S1, S2, systole, diastole respectively). We also implemented the parameter, which describes value of the main peak frequency divided by the width of the peaks for S1 intervals. As the next parameter we considered width of the main peak in the S1 interval. Finally we added to the features set the average powers in the respective frequency ranges (0-50 Hz or 100-200Hz) for S1, systole, S2 and diastole phases.

All of the features contributed to the network by indicating whether the rhythmicity in the signal was maintained.

2.4 Neural network characterization and training

We used two types of neural network architectures: standard multilayer perceptron architecture with 1 hidden layer and deep network architecture with 2 hidden layers. In both cases the input layer size was equivalent to the number of features (48 features for the best score) and the output layer was 1 neuron with linear transfer function giving a classification result (1 for positive sample or -1 for negative sample).

We determined the hidden layer size in the standard architecture by plotting the optimization surface presenting the classification error (Mean Squared Error) on the validation subset in the range of 31–71 hidden neurons (sigmoid transfer function). The best architecture was 59 hidden neurons and training time of 130 epochs (for Scaled Conjugate Gradient training algorithm) or 35 epochs (for Levenberg–Marquardt training algorithm). We also used the regularization parameter of 0.1 to prevent overfitting on the training subset. The number of 59 neurons was adopted as the first hidden layer in the deep network.

In the case of the deep network architecture we performed pre-training of the hidden layers (so called autoencoders). The pre-training involved training each

layer in a such way that the output of the layer was the same as the input. Therefore the weights in the layer are closer to the proper place in the weight space and final training can be effective. When the size of the feature vector in the autoencoder is greater than the size of the layer, the pre-training causes that the information from the feature vector is compressed to the smaller number of neurons. At the final stage all the hidden layers (59 and 15 neurons) and the output layer were stacked into one network and the final training (fine tuning) was performed.

The original output of the linear neuron in the output layer was linear function ranging from 0 to 1 which was mapped to -1 and 1 with the threshold of 0.5. We set this threshold to 0.75 (increasing specificity).

3. Results

For the training set, our algorithm performed with sensitivity of 0.83 and specificity of 0.62, giving overall score equal to 0.73. For the test set, our final result in PhysioNet Challenge 2016 equaled 0.79. The algorithm performed with sensitivity of 0.81 and specificity of 0.76.

4. Discussion & conclusions

The proposed algorithm consisted of three modules: assessment of signal quality, pre-detection of abnormal signals and neural networks.

The advantage of direct classification of the poor quality signals was that the features provided later to the neural network came only from proper PCG signals and supplied it with crucial information about the said signal. The pre-detection module was characterized by very high precision, what kept us from increasing the false positive classifications. But it's extremely low specificity prevented the algorithm from detecting more complex abnormal recordings, what should be improved to arise algorithm efficiency. As for now, this method is extremely effective in detecting signals from patients with mitral regurgitation, which represented the majority of recordings classified as abnormal in this step.

Current algorithm does not have any mechanism to cope with artifacts occurring irregularly, due to external factors e.g. opening and closing of the doors, machine hum, human voices or technical problems such as moving the chest piece of stethoscope on the skin during measurement (In cases where the level of such disturbance is extremely high it is detected by the Signal Quality Assessment method – see fig. 1.). Now, we eliminate influence of such interruptions on our features by taking the statistical majority of the results leaving the margins to mentioned interruptions. The appropriate preprocessing of signal might assure higher quality of latter analysis and classification.

The performance of the neural network part of the algorithm was sufficient, but we believe that the results would be significantly better if the information about the location of the stethoscope on the chest was more detailed. Then, we would be able to adapt the features for each location and look for the particular pathology. What is more, supporting the choice of the final features to the network by using another learning algorithm might prove to be more accurate.

References

- [1] Liu C Y, Springer D B, Li Q et al. 2016 An open access database for the evaluation of heart sound algorithms *Physiol Meas* 37 (in press).
- [2] Schmidt, S.E.; Holst-Hansen, C.; Graff, C. et al. Segmentation of heart sound recordings by a duration-dependent hidden markov model. *Physiol Meas* 2010, 31, 513-529.
- [3] Springer, D.B.; Tarassenko, L.; Clifford, G.D. Logistic regression-hsmm-based heart sound segmentation. *IEEE Trans Biomed Eng* 2015, In press.
- [4] Behar Joachim, Jonhson Alistair, Clifford Gari D., Oster Julien A Comparison of Single Channel Fetal ECG Extraction Methods. *Annals of Biomedical Engineering* 42(6), 1340-53. 2014
- [5] Gierałowski, J., Ciuchciński, K., Grzegorzczak, I. et al. (2015). RS slope detection algorithm for extraction of heart rate from noisy, multimodal recordings. *Physiological measurement*, 36(8), 1743.
- [6] Gupta, C. N., Palaniappan, R., Rajan et al. (2005, May). Segmentation and classification of heart sounds. In *Canadian Conference on Electrical and Computer Engineering*, 2005. (pp. 1674-1677). IEEE.

Address for correspondence:

Iga Grzegorzczak,
Warsaw University of Technology, Faculty of Physics
Koszykowa 75 00-662 Warsaw, Poland
igagrze@gmail.com