Machine Learning Methods in Electrocardiography Classification

Iryna Mykoliuk[1], Daniel Jancarczyk[1], Mikolaj Karpinski[1], Viktor Kifer[2]

1. Department of Computer Science and Automatics, University of Bielsko-Biala, POLAND, Bielsko-Biala, 2 Willowa St, email: i.mykoliuk@gmail.com, djancarczyk@ath.bielsko.pl, mkarpinski@ath.bielsko.pl

2. Faculty of Computer Information Systems and Software Engineering, Ternopil Ivan Puluj National Technical University, UKRAINE, Ternopil, 56 Ruska St, email: kifervictor@gmail.com

Abstract. **This document reports results that tend to confirm the applicability of the machine learning combined with signal processing for automatic atrial fibrillation detection from a short single lead electrocardiography recording.**

Keywords: **machine learning, signal, ECG, classification, heartbeat.**

I. Introduction

Automated electrocardiography (ECG) analysis has a number of ground tasks which include noise removal, QRS detection, P and T waves detection etc. The first two problems already have a number of techniques which provide good results. In particular band pass filters, Fourier based analysis and transforms and wavelet transformations are commonly used for noise reduction. Pan Tompkins algorithm, various transforms like Wavelet, Hilbert and Empirical Mode Decomposition combined with some decision logic can identify QRS complexes with detection rates over 99%.

P and T wave have small amplitudes, which makes them less noise resistant. But position and shape of the waves are the important components of ECG analysis. With the most recent approaches based on advanced Kalman filters and wavelet transforms, detection accuracy for those waves reached over 90%.

In general, ECG classification solutions tend to be class testing, when a particular record is checked against some abnormal rhythm pattern. This happens because there's a large variety of ECG shape changes, which are hard to handle by a single algorithm. And there is a strong tendency in last decade of moving from a threshold-based analysis (which suffers from introducing to validation dataset ECG recordings from new patients or recorded with new devices) to data-driven approaches which include classical supervised machine-learning models and neural networks.

II. Biological background

Understanding of heart activity is an important part. Human heart consists of four chambers: left atrium (LA), right atrium (RA), left ventricle (LV) and right ventricle (RV). Between RA and RV there is a tricuspid valve, and between LA and LV there is a mitral valve. These valves prevent blood going directly from atriums to ventricles. Ventricles also have valves which prevent blood from going into veins, there is a pulmonary valve in right ventricle and an aortic valve in the left one.

Electrocardiogram

ECG is the process of recording the electrical activity of the heart over a period of time using electrodes placed on the skin. These electrodes detect the tiny electrical changes on the skin that arises from the heart muscles.

In a conventional 12-lead ECG, 6 electrodes are placed on the surface of the chest, 2 on hands, and 2 on legs. The overall magnitude of the heart's electrical potential is then measured from 12 different angles ("leads") and is recorded over a period of time. A healthy heart has a specific order of polarization and depolarization during each heartbeat. It starts with the sinoatrial node, then spreads through the atrium to atrioventricular node, and then to ventricles.

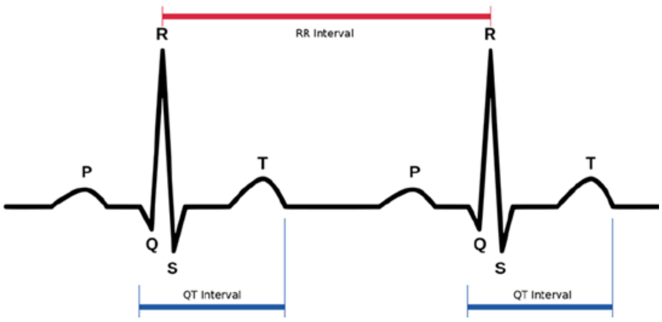

Fig. 1.Typical ECG trace

The graph of voltage versus time produced by this noninvasive procedure is called electrocardiogram. Each ECG consists of periodic PQRST complexes which represent a cycle of heart activity. In particular, one complex has:

1. P-wave (atrial contraction)
2. QRS-complex (contraction of the ventricles)
3. T-wave (relaxation of the ventricles)

The duration of PQRST might differ depending on the heart rate. There are a couple of ways how to measure it, one of the most common is called RR-interval, which represents the duration as a distance between two following R peaks

Common abnormal heart rhythms

Heart irregular rhythm can be classified by heart rate into tachycardia (heart beats too fast, more than 90 bpm) and bradycardia (heart beats too slow, less than 60 bpm); and by place of occurrence into supraventricular (atria contracts irregularly) and ventricular (ventricles contract in an irregular pattern) arrhythmias.

Atrial fibrillation is an abnormal heart rhythm characterized by rapid and irregular beating and which has no symptoms in most of the cases. It is the most common serious abnormal heart rhythm and it affects approximately 2-3% of the population in Europe and North America. Also, the percentage of people with AF increases with age with 0.14% under 50 years old, 4% between 60 and 70 years old and 14% over 80 years old being affected. On ECG, atrial fibrillation is usually diagnosed by absence of P-wave and irregular heart beats pattern.

Atrial flutter (AFL) is an abnormal heart rhythm, similar to atrial fibrillation. Both of them are types of supraventricular tachycardia. In AFL the electrical signal goes along the pathway in a circular motion, which results in atrial muscle contractions that are faster than and out of sync with the lower chambers (ventricles). Similar to A-Fib the heart beats fast, but in a regular pattern.

Itself atrial flutter is not life-threatening, but due to slower blood pumping it creates, the side effects might cause health problems.

On ECG atrial flutter might be diagnosed by the presence of multiple f-waves instead of a P-wave.

Paroxysmal supraventricular tachycardia (PSVT) is characterized by the episodes of a rapid increase of heart rate (100-250 bpm) and usually regular heart rhythm. It is most often seen in young people and infants. Alcohol, illicit drugs, caffeine, and smoking might be a cause of this arrhythmia.

Another abnormal heart rhythm is premature ventricular complex (PVC) characterized by long QRS complexes (> 0.12ms), ST segment and T wave changes (irregular, inverted in leads with a dominant R wave, etc).

Ventricular fibrillation (V-Fib) is a life-threatening abnormal heart rhythm when the lower chambers quiver and the heart can't pump blood to the body. This occurs because of electrical signal disorder which makes chambers contract very fast in an unsynchronized way. It requires immediate treatment and can usually be stopped with a defibrillator.

III. Data preprocessing

Handling of imbalanced data

The data in selected dataset is quite imbalanced which might create problems during training. The model might easily overfit and predict all of the time the most represented category. This creates a so-called accuracy paradox: suppose that we have a dataset with two labels A and B and 90% of the records are instances of A, the classifier might predict all the time class A which will give the accuracy of 90% although it definitely won't be a good classifier. There are several common approaches here: use abnormalities detection algorithms or apply class weights or perform data balancing.

The first approach was applying class weights as a logarithmic function of the proportion of total instances divided by the number of class instances.

$$classweight = min\left(log\left(\frac{\mu * totalinstances}{classinstances}\right), 1\right)$$

where μ - parameter to tune

While applying class weight helped to prevent overfitting on early training stages, the model still failed to learn class-specific features.

The second approach was to apply data balancing. There are many techniques available for performing data balancing including:

Under-sampling (deleting instances of over-represented classes)

Over-sampling (repeating training with under-represented classes)

Generating Synthetic Samples (generating new samples of the under-represented based on available samples)

Use algorithms which are better in handling imbalanced data.

Under-sampling, over-sampling and using of decision trees which can handle imbalanced data were used to work out the problem.

Normalizing data

Measuring ECG values might generate a wide range of values depending on different conditions when the recording was performed. This might misdirect model to learn absolute values of one instance instead of the value for all instances.

To prevent this situation, data normalization is applied which means adjusting measured values to a common scale. The first step is to find a baseline (mean value of the record) and subtract it from the record values. And after that divide record values by the absolute value of the record. So the final scale would be in the range [-1; 1].

Band pass filter

ECG signal might be altered with the noise of different sources. This might be electrical signals coming from human body activity (usually electromyography signals from muscles contraction), procedure artifacts (eg. connecting/ disconnecting electrodes), power line interference and instrumentation noise.

Power line interference is not less represented in hand-held devices, and usually, occurs through inductive mechanism. The power lines across the world have the frequency of 50Hz or 60Hz, which is higher than the subject information frequencies (0.1Hz to 40Hz). So it can easily be eliminated by filters as presented on Fig. 2.

Both electrode contact noise and electrode motion noise cause baseline changes that occur due to variation in the position of the heart with the respect to the electrodes. More specifically, the amplitude of the changes is defined by variation of electrode-skin impedance (ESIV). The larger ESIV, the smaller change of impedance needed to cause a shift in the baseline. This kind of noise is not defined by any particular range of frequencies and it isn't easy to eliminate it with signal processing, the task is to perform the recording procedure without patient disturbance.

Electromyography (EMG) noise is caused by contraction of muscles, not related to the heart activity (moving of hands and/or legs). For the hand-held devices usually, the highest EMG noise component is waves generated by depolarization and repolarization of hand muscles (as ECG signal goes from heart to the device directly through hands). EMG signal frequency is stochastic in nature, but the significant activity happens in the range of 5Hz and 450Hz. This partially overlaps with the frequencies of the ECG signal so it's hard to eliminate it absolutely. In particular, the problematic areas are P and T waves of the ECG.

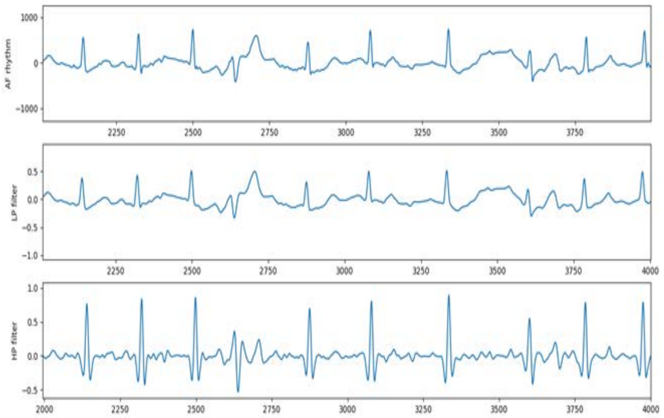

Fig.2. Signal preprocessing with low-pass and high-pass band filters

So to remove most of the noise the combination of low pass and high pass bands filters were used. The range of frequencies to be kept is from 0.1Hz to 40Hz.

The SciPy filter function in combination with NumPy convolve function was used to implement both filters.

IV. Feature selection

Feature selection includes:

1. QRS detection.
2. Heartbeat extraction.
3. Features extraction

The task of QRS detection is to find R-peaks which represent the heart beats. Beat detection is a procedure preceding any kind of ECG analysis and so it is critical to correctly detect heart beats.

One of the standard algorithms used for QRS detection is a Pan-Tompkins algorithm. The original paper states detection rate of 99.3%. Many modifications were made to the algorithm since its publication including additional filtering, search back support, outliers rejection etc.

The filtered signal is introduced as an input to the algorithm. The processing of the data starts with a first-derivative filter, which helps obtain information about the slope of the QRS.

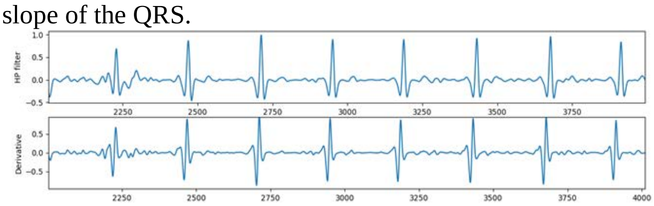

Fig.3. Signal before and after derivative filter

To get even better intensity of the slope, the squaring with normalization is applied. This gives a graph where most of the values are near to zero, and in the positions of R peaks, it has a value much higher (usually over 0.5).

Additionally moving average is applied, which combines two close peaks on the graph above into one. After that, the algorithm finds positions on the graph that are higher than the average (referenced threshold). And for all of those positions it looks for the position of the higher value. This position is marked as potential R-peak location.

R-peaks than go through false rejection algorithm which checks the average length of the RR-interval, and rejects those R-peaks, where the length of RR-interval is less than half of the average.

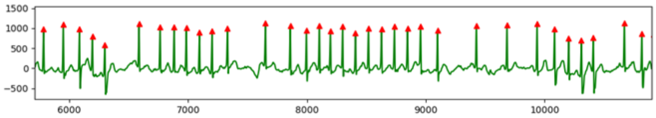

Fig.4. Detection of R-peaks

There is also additional search back algorithm which again looks on RR-intervals and finds positions where there might be a missed R-peak (RR-interval is higher than 1.5 of the average). There it lowers the threshold and runs the search algorithm once again.

Heartbeat extraction is done with the static window. By taking in consideration that normal heart rhythm is in the range of 60 to 100 beats per minute, we considered static windows of sizes from 0.5s (120 bpm) to 0.7s (86 bpm). The window size that served best was 0.6s with the distribution of 0.15s for P-wave, 0.1s for QRS complex and 0.35s for ST interval and T-wave.

For each R-peak detected, the heartbeat template was extracted with the static window of 0.6s which includes 0.2s before and 0.4s after R-peak.

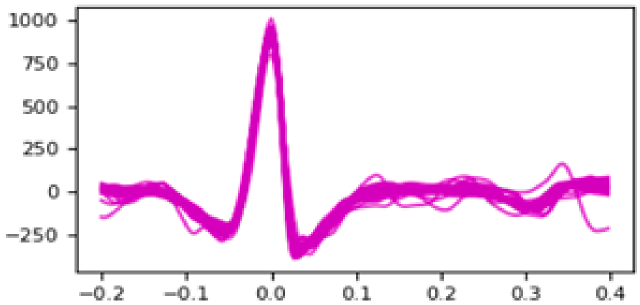

Fig. 3. Extracted heartbeat

During the model development process, many features were considered as good for prediction. To analyze the noise of the data, transformation coefficients were used. In particular, first 400 coefficients of Fourier transformation and details coefficients of 4-level wavelet decomposition.

Another range of useful features is heart rate variability. It is computed on RR-intervals, which includes time domain components (RMSSD, SDNN, NN20, PNN20, NN50, PNN50, mean RR, std RR, mean HR) and frequency domain components.

Frequency domain components are computed for 3 frequency bands: very low frequencies (0-0.04), low frequencies (0.04-0.15) and high frequencies (0.15-0.4). To compute frequencies powers, Welch estimation of power spectral density is performed on interpolation of RR-intervals, and the total values of power bands are computed by integrating (with composite trapezoidal rule) function of frequencies and powers.

Additionally, minimum and maximum RR-interval lengths are added to the feature vector.

The extractor also computes the proportion of the number of R-peaks detected divided by the length of the record, mean and standard deviation of normalized R-peak values, number, and proportion of detected R-peak values with the inverse sign.

Based on the templates of heart beats the module finds a median heartbeat as it best represents the most common template of heartbeat. Median was used rather than mean because it is more noise resistant. The extractor also computes the standard deviation of each point of templates.

V. Classification

One of approaches to ECG classification was based on decision trees. The bagging was used from the very beginning with the random forest algorithm.

Random forest implies divide-and-conquer principle, which makes it easy to execute in parallel. This makes training process very fast (a couple of seconds). This is particularly good for experiments with the trees structures optimizations. For classification process was selected Random forest algorithm which consists of 60 decision trees. The implementation of the algorithm is taken from the Scikit-learn library. Results of classification process for random forest based model shown in table 1.

TABLE 1. Scores for Random forest model.

Class identifier	Precision	Recall	f1-score
0	0.77	0.67	0.72
1	0.73	0.75	0.74
2	0.64	0.68	0.66
3	0.68	0.61	0.65
Avg/total	0.70	0.70	0.70

Another approach was based on fully convolutional network, this type of network architecture serves as a feature extractor, and the final decision is made by a fully-connected layer. Results of classification process for neural network based model shown in table 2.

TABLE 2. Scores for neural network model.

Class identifier	Precision	Recall	f1-score
0	0.58	0.26	0.36
1	0.65	0.94	0.77
2	0.60	0.28	0.38
3	0.47	0.03	0.06
Avg/total	0.62	0.64	0.59

Model used 4 categories '0' – is atrial fibrillation, '1' – is normal sinus rhythm, '2' – is other heart rhythm and '3' – is noisy signal. Precision is defined as the number of true positives divided by the number of true positives and false positives and is a measure of a class exactness. Recall is the number of true positives divided by the sum of a number of true positives and false negatives and it expresses the measure of classifier completeness. F1 score is used as a model estimate. F1 measure is defined a combination of specificity and sensitivity:

$$F1 = \frac{2 * precision * recall}{Precision + recall}$$

It can be used to express the average of classifier exactness and completeness. Scoring algorithm include F1 metrics only for 3 categories: atrial fibrillation, normal rhythm, and other rhythm:

$$F1 = \frac{F1(AF) + F1(N) + F1(O)}{3}$$

So the final score of the model is an average of F1 scores for 3 classes.

VI. Conclusions

This paper shows the ECG automated analysis problem and one of suitable algorithms for this – Random forest. Was described the main steps of ECG signal preprocessing which includes handling of imbalanced data, data normalization and noises filtering. Also presented QRS complexes detection and the main features for ECG classification.

References

[1] Heart Disease Facts // https://www.cdc.gov/ URL: https://www.cdc.gov/heartdisease/facts.htm (last accessed: 2018/03/28).

[2] Wang, Zh., Yan, W., Oates, T. Time Series Classification from Scratch with Deep Neural Networks: A Strong Baseline *// arXiv preprint. arXiv:*1611.06455, 2016.

[3] Brownlee J. Classification accuracy is not enough: More performance measures you can use // http://machinelearningmastery.com/ URL: http://machinelearningmastery.com/classification-accuracy-is-not-enough-more-performance-measures-you-can-use (last accessed: 2018/03/28).

[4] ECG basics // https://www.osmosis.org/ URL: https://www.osmosis.org/learn/ECG_basics (last accessed: 2018/03/28).

[5] Johnson, A. E., Behar, J., Andreotti, F., Clifford, G. D., & Oster, J. (2014, September). R-peak estimation using multimodal lead switching. *In Computing in Cardiology Conference (CinC)*, 2014 (pp. 281-284). IEEE. Thomas Mitchell. Machine Learning / Thomas Mitchell. – *McGraw-Hill*, New York, 1997. – ISBN:0070428077.