New Indices for Sleep Apnea Detection from Long-Time ECG Recordings

Agata Pietrzak[1], Gerard Cybulski[1,2]

[1]Institute of Metrology and Biomedical Engineering, Department of Mechatronics, Warsaw University of Technology, Warsaw, Poland
[2]Mossakowski Medical Research Centre, Polish Academy of Sciences, Warsaw, Poland

Abstract

We used our computer program enabling detection of sleep apnea using long-time one-channel ECG signal recordings. It allows the calculations of commonly accepted six heart rate variability (HRV) parameters in time domain. We also introduced additional 34 indices which were created as a combination of selected or all basic six indices of HRV. For testing we used 70 sample recordings from the Physionet database containing single ECG signals 7 to 10 hours long. The analysis was performed on samples lasting 10000 seconds. The efficiency of the software was evaluated using the Receiver Operating Characteristic (ROC) method. For basic 6 HRV indices we found that the highest accuracy of discrimination was achieved for standard deviation of successive differences (88.5%). The area under the ROC curve was 0.89. The sensitivity and specifity were 96% and 70%, respectively. For one of the newly proposed indices which was average sum of square of all six base indices the accuracy was at the level of 90%. The area under the ROC curve was 0.85. The sensitivity and specifity were 98% and 70%, respectively.

1. Introduction

Sleep apnea is a disorder characterized by abnormal pauses in breathing or episodes of abnormally low breathing during sleep. Pauses can last from a few seconds to a few minutes, and may occur 5 to 30 or more times per hour [1]. The symptoms of sleep apnea may include morning headaches, memory or learning problems, difficulty concentrating, unstable emotional states (irritation, depression, or mood swings), or urination at night. Untreated sleep apnea, apart from causing discomfort, can lead to the destruction of parts of the central nervous system and/or increase the risk of high blood pressure, heart attack, stroke, obesity, and diabetes. It may increase the risk of heart failure and arrhythmias. It may also increases the chance of having work-related accidents or driving accidents due to falling asleep at the wheel. Unfortunately, patients with sleep apnea are rarely aware of the breathing difficulty, even upon awakening [1], [2], [3]. Sleep apnea can be recognized by others witnessing episodes or suspected because of its effects on the body. Symptoms may be present for years without identification. Over this period, the sufferer may become conditioned to the daytime sleepiness and fatigue associated with significant levels of sleep disturbance [1], [2]. In 1993, a cohort study estimated that roughly one in every 15 Americans was affected by at least moderate sleep apnea [1], [2]. It also estimated that in middle-age as many as 24 percent of men and 9 percent of women and were affected, undiagnosed, and untreated [1], [2], [3]. Thus, the early detection of sleep apnea occurrence is an important clinical challenge. Although polysomnography is the gold standard in detection of all types of sleep apnea, the search continues for a simple method for detecting these episodes based on a limited source of information (e.g. single ECG recording). There are several methods of detecting the sleep apnea symptoms from ECG signals. They can be divided into two groups. Methods in the first group are based on detection of the respiratory signal from ECG [4], [5] and identification of the moment of a sleep apnea episode's occurrence. In the second group are methods based on extraction of ECG signal features which can be used to detect if episodes occurred without determining their timing [6], [7].

The aim of this paper was to develop computer software enabling the analysis of long-time, one-channel ECG recordings. It was also intended to check if some heart rate variability parameters could be used to create a classification model based on the Support Vector Machines (SVM) method using the discriminative Radial Basis Function (RBF) kernel.

2. Material and methods

In this study we used our computer software enabling detection of sleep apnea using long-time one-channel ECG signal recordings. It allows the calculations of commonly accepted six heart rate variability (HRV) parameters in time domain. We also introduced additional 34 indices which were created as a combination of selected or all basic six indices of HRV.

2.1. Computer software

The computer program for detection of sleep apnea using a one-channel ECG signal - Detection of Obstructive Sleep Apnea (DOOSA) with Graphical User Interface was prepared in the MatLab environment (v. 2013a). The program consists of four parts. In the first, data acquisition and preparation is performed. The second part contains procedures of QRS complex detection with RR-interval calculation based on the Pan-Tompkins method [8]. The third component of the program allows extraction of basic heart rate variability (HRV) parameters. The last part performs calculations associated with application of the Support Vector Machine (SVM) method [9], [10], for classification of the signals.

2.2. Database and data acquisition

Data were imported from the Physionet webpage's "Apnea-ECG Database (apnea-ecg)" [11] in European Data Format (EDF) [12]. The database contains 70 recordings of one-channel ECG lasting 7-10 hours. Data were recorded with a sampling frequency of 100Hz and a resolution of 16 bits, where 1 bit denotes 5μV. The analysed signals were arranged in records of 2000, 4000, 6000, 8000, and 10,000 seconds. The signals were classified by providers according to the Apnea–Hypopnea Index (AHI), which represents the number of apnea and hypopnea events per hour of sleep. There are four categories of signals according to AHI: normal (23 records), when AHI is between 0 and 4, mild sleep apnea (3 records, AHI=5…14), moderate sleep apnea (13 records, AHI=15…30), and severe sleep apnea (31 records, AHI>30).

2.3. QRS detection and RR interval series

The procedure for QRS complex detection was based on Pan-Tompkins algorithm [8]. It contains the following stages: band pass filtering, differentiation, exponentiation, moving window integration, and thresholding. When QRS complexes are identified (R-waves marked), the matrix of RR intervals is created and stored.

2.4. Heart rate variability parameters

Based on the RR-interval lengths series, the commonly accepted heart rate variability parameters in time domain are calculated: the mean value of the RR interval (MeanNN), the standard deviation of the successive differences between adjacent NN's (SDNN), the root mean square of successive differences (RMSSD), the number of pairs of successive NN's that differ by more than 50 ms (NN50), the proportion of NN50 divided by total number of NN's (pNN50), and the standard deviation of successive differences (SDSD). Their definitions are presented below according to the recommendations [13]:

$$MeanNN = \overline{RR} = \frac{1}{N}\sum_{i=1}^{N} RR_i, \tag{1}$$

$$SDNN = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(\overline{RR} - RR_i)^2}, \tag{2}$$

$$RMSSD = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N-1}(RR_{i+1} - RR_i)^2}, \tag{3}$$

$$NN50 = \sum_{i=1}^{N-1} f_i, \tag{4}$$

$$pNN50 = \frac{NN50}{N-1} \times 100\%, \tag{5}$$

$$SDSD = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N-1}(RR_{i+1} - RR_i)^2}, \tag{6}$$

where $$f_i = \begin{cases} 1\ gdy\ (RR_{i+1} - RR_i) > 50ms \\ 0\ gdy\ (RR_{i+1} - RR_i) \le 50ms \end{cases}, \tag{7}$$

N = number of records and RR_i = value of ith RR interval length in [ms].

2.5. New indices of HRV

We introduced additional 34 indices which were created as a combination of selected or all basic six indices of HRV. They are created as product, sum of squares and ratios of all or selected basic indices. Some of them are expressed as normalised values.

2.6. Classification

With the aim of identifying sleep apnea using HRV parameters, the Support Vector Machine (SVM) method, described by Vapnik [9], was used. The program used the procedure developed by Kris De Brabanter et al. [10] based on Least Squares SVM using the discriminative Gaussian RBF as a kernel function in the learning process for separation of "occurrence" from "normal activity" for two sets of data: learning and testing.

The Gaussian RBF kernel describes the boundary between classes separated nonlinearly. It is determined using the following formula:

$$\mathrm{K}(\mathrm{xi},\mathrm{xj}) = e^{-\frac{||\mathrm{xi}-\mathrm{xj}||^2}{2\sigma^2}}, \tag{8}$$

where $||\mathrm{xi} - \mathrm{xj}||^2$ represents the squared Euclidean distance between the two feature vectors and σ is a parameter.

The leaning process was performed on 35 of 70 signals (lasting 10000 seconds) from database [11] with "non-apnea” coded as -1 and “apnea” as 1. The learning set was arranged as a matrix of 35 rows (number of learning signals) by 6 columns (number of analyzed HRV parameters described by formulas: 1…6). Another matrix (vector) of 35 rows (number of learning signals) by 1 column contained the “non-apnea”/“apnea” codes. The test signals (another 35 from the set of 70, lasting 10000 seconds) were arranged similarly to the learning signals.

The efficiency of the obstructive sleep apnea detection software was evaluated using the Receiver Operating Characteristic (ROC) method by comparing the automatically derived results with true results. The program draws the ROC curve and calculates the area under the curve (AUC) and the cost-effective cut-off point (CUT). The AUC is equal to the probability that a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative one (assuming “positive” ranks higher than “negative”). The CUT is the point on the ROC curve the shortest distance from the upper left corner. It corresponds to the optimal cut-off value for the form factor. Assuming the cut-off point’s optimal level, the confusion matrix containing the true positive (TP), true negative (TN), false positive (FP), and false negative (FN) was determined. Also calculated were the:

- sensitivity – SENS = TP/(TP+FN),
- specificity – SPEC = TN/(TN+FP),
- positive predictive value – PPV = TP/(TP+FP),
- negative predictive value – NPV= TN/(TN+FN),
- accuracy – ACC = (TP+TN)/(TP+TN+FP+FN).

3. Results

Table 1 contains the averages of the heart rate variability parameters calculated according to formulas (1)…(6) in subjects with no sleep apnea and for all apnea groups together (with mild, moderate, severe apnea) based on 10000 seconds of recording.

Table 2 contains parameters which describe classification quality: area under the curve (AUC) from Receiver Operating Characteristic analysis and accuracy (ACC), as a measure of classification efficiency. AUC and ACC are expressed as fraction and have no units (value between 0 and 1).

The highest sensitivity was observed for the MeanRR index (97%) and the highest specificity was achieved for the pNN50 index (83%). Since the specificity of MeanRR was very low (43%, the minimal for analyzed indices), the overall accuracy was at the level of 80%. For MeanNN it was also observed the highest negative predictive value (91%), whereas the positive predictive value was 77%.

For one of the newly proposed indices which was average sum of square of all six base indices the accuracy was at the level of 90%. The area under the ROC curve was 0.85. The sensitivity and specifity were 98% and 70%, respectively. The positive and negative predicting value were 0.87 and 0.94, respectively.

Table 1. Average values of basic heart rate variability (HRV) in healthy and sleep apnea subjects.

HRV parameter	Healthy	Apnea
MeanNN [ms]	941	875
SDNN [ms]	103	376
RMSSD [ms]	103	373
NN50	2170	1277
pNN50 [%]	49	15
SDSD [ms]	103	346

Table 2. Accuracy classification parameter (ACC) and area under the curve (AUC) from ROC analysis.

HRV parameter	AUC	ACC
MeanNN [ms]	0.70	0.81
SDNN [ms]	0.88	0.87
RMSSD [ms]	0.88	0.87
NN50	0.55	0.69
pNN50 [%]	0.89	0.87
SDSD [ms]	0.88	0.89

Fig. 1 presents the ROC curve for the newly introduced index with the best accuracy (average sum of square of all six base indices).

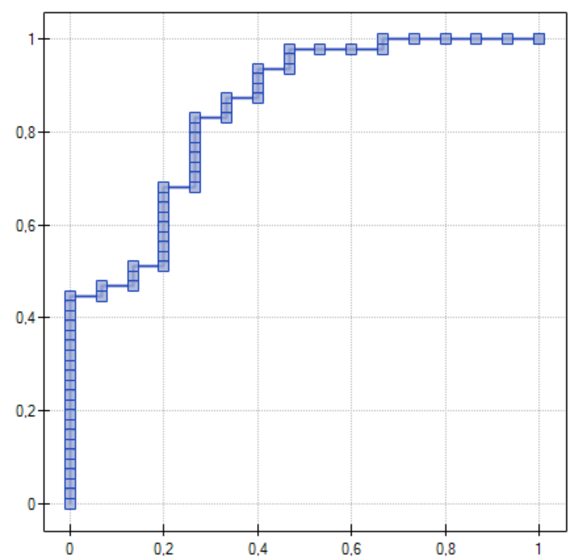

Figure 1. ROC curve for the average sum of square of all six base indices. Sensitivity and 1-Specifity are on vertical and horizontal axis, respectively.

For SVM classification when all six basic indices (formulas: 1-6) were used as an input ACC was 0.914 and AUC - 0.931.

4. Discussion and conclusions

A tool was developed enabling calculation of heart rate variability parameters and their use as input data for a sleep apnea classifier based on a Support Vector Machine using the discriminative Gaussian RBF-kernel function. The idea of developing computer software which can help in off-line analysis of numerous overnight ambulatory ECG recordings with the aim of detecting sleep apnea's presence seems very attractive from a diagnostic point of view. Even the calculation of simple HRV parameters can notify that there are different values for patients compared with healthy subjects, but there is no linear relationship between the AHI and any of the tested parameters. Thus, a more advanced classification tool is essential to detect sleep apnea occurrence. The preliminary results suggests that DOOSA software classifies signals as 'with apnea' and 'without apnea' with 80% accuracy, which is in the lower part of the range found in the literature [4], [14], [15]. We are not fully satisfied with the results obtained using new indices. Also operation on basic indices did not revealed satisfactory increase in sleep apnea detecting efficiency. Perhaps the search for new approach is essential. One of the study limitations regarded a relatively small amount of data the analysis was performed on. However, it seems that the tool might be used as a base for further development of sleep apnea detection using ECG signals.

Acknowledgements

Study supported by research programs of institutions the authors are affiliated with.

References

[1] Lee W, Nagubadi S, Kryger MH, Mokhlesi B. Epidemiology of obstructive sleep apnea: a population-based perspective. Expert Rev Respir Med., 2008, 2 (3): 349–64.
[2] Young T, Palta M, Dempsey J, Skatrud J, Weber S, Badr S. The occurrence of sleep-disordered breathing among middle-aged adults. The New England Journal of Medicine, 1993, 328 (17), pp. 1230–5.
[3] Young T, Peppard PE, Gottlieb DJ. Epidemiology of obstructive sleep apnea: a population health perspective. American Journal of Respiratory and Critical Care Medicine, 2002, 165 (9):1217–39.
[4] Chazal P, T. Penzel T, Heneghan C. Automated Detection of Obstructive Sleep Apnoea at Different Time Scales Using the Electrocardiogram, Physiol Meas. 2004, 25(4): 967-83.
[5] Yilmaz B, Asyali M, Arikan E, Yektin S, Ozgen F. Sleep Stage and Obstructive Apneaic Epoch Classification Using Single-Lead ECG, in Biomedical Engineering Online, 2010. vol. 9: 39
[6] Khandoker AH, Palaniswami M, Karmakar C. Support vector machines for automated recognition of obstructive sleep apnoea syndrome from ECG recordings, IEEE Trans. Inf. Technol. Biomed., 2009, vol. 13(1):37–48.
[7] Almazaideh L, Elleithy K, Faezipour M. Obstructive sleep apnoea detection using SVM-based classification of ECG signal features", Proc. of the Annual International Conference of the IEEE EMBS, 2012:4938-41.
[8] Pan J, Tompkins WJ. A Real-Time QRS Detection Algorithm, IEEE Transactions on Biomedical Engineering, 1985, VOL. BME-32, 3:230-36.
[9] Cortes C, Vapnik V. Support-Vector Networks. Machine Learning, 1995, Vol. 20, 3:273-97.
[10] De Brabanter K, De Brabanter J, Suykens JAK, De Moor B. Approximate Confidence and Prediction Intervals for Least Squares Support Vector Regression, IEEE Transactions on Neural Networks, 2011, vol. 22, no. 1, pp. 110-20.
[11] http://www.physionet.org/physiobank/database/apnea-ecg
[12] http://www.edfplus.info
[13] Heart Rate Variability: Standards of Measurement, Physiological Interpretation, and Clinical Use, Task Force of the European Society of Cardiology, the North American Society of Pacing Electrophysiology, Circulation, 1996;93: 1043-65.
[14] Manrique Q, Hernandez A, Gonzalez T, Pallester F, and Dominquez C. Detection of Obstructive Sleep Apnea in ECG Recordings Using Time-Frequency Distributions and Dynamic Features, in Proceedings of the IEEE International Conference on Engineering in Medicine and Biology Society(EMBS 2009), pp. 5559-62.
[15] M. Mendez, D. Ruini, O. Villantieri, M. Matteucci, T. Penzel and A. Bianchi. Detection of Sleep Apnea from Surface ECG Based on Features Extracted by an Autoregressive Model, in Proceedings of the IEEE International Conference on Engineering in Medicine and Biology Society (EMBS 2007):6105-8.

Address for correspondence.

Gerard Cybulski
Warsaw University of Technology
Faculty of Mechatronics
Institute of Metrology and Biomedical Engineering
Sw. A. Boboli 8,
02-575 Warsaw
Poland
G.Cybulski@mchtr.pw.edu.pl