

Sleep Apnea Detection Using Multi-Lag Poincare Plot

Shahab Rezaei¹, Sadaf Moharreri², Nader Jafarnia Dabanloo³, Keivan Maghooli³, Saman Parvaneh^{3, 4}

¹ Department of Biomedical Engineering, Central Tehran Branch, Islamic Azad University, Tehran, Iran

² Department of Biomedical Engineering, Khomeini Shahr Branch, Islamic Azad University, Isfahan, Iran

³ Department of Biomedical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran

⁴ Edwards Lifesciences, Irvine, CA, USA

Abstract

The Poincare plot is a geometrical representation of RR time series to study heart rate dynamics. Analysis of Poincare plot with different lags provides insights into the autonomic control of the heart. In this research, various descriptors of points in the Poincare plot of RR time series reconstructed in 10 different lags are used for detecting sleep apnea. The Poincare plot of RR time series was reconstructed using ten different lags (1-10). Traditional ($SD1$, $SD2$, $SD1/SD2$, and $SDNN$) and advanced features (angle and the average of point's distances relative to the line of identity) were extracted for each lag. Four groups of 70 ECG recordings (three classes in the Apnea-ECG database and a class in the Normal Sinus Rhythm database from the Physionet database) were used to evaluate the new features and their application for apnea detection. The duration of all the recordings was 5 minutes. The classifier was trained on 80% of data as a train set, and the accuracy was evaluated on 20% of data as a test set. In the classification of three apnea classes from Normal Sinus Rhythm using six extracted features from Poincare plot in lag equal to one, random forest classifier led to the best AUC equal to 0.889 and 0.87 for 5-folds cross-validation and test set. The AUC of 0.906 and 0.87 were obtained by the XGBoost model using 60 extracted features from the Poincare plot in lag one to ten in 5-folds cross-validation and test set, respectively.

1. Introduction

Sleep apnea is a sleep disorder that causes breathing stops during sleep [1]. Each pause can last for a few seconds and is usually accompanied by a loud snoring sound and sometimes causes the person to suddenly wake up suffocated. During apnea, there is a sudden drop in blood oxygen levels [2]. This raises blood pressure and puts pressure on the cardiovascular system. Obstructive sleep apnea can increase the risk of heart attack, stroke, and abnormal heartbeats such as atrial fibrillation [3].

Therefore, the need for timely and rapid diagnosis of apnea has attracted numerous studies' attention. Considering the effects of apnea on the activity of the cardiovascular system, one of the ways to diagnose this complication is to study the behavior of heart rate variability (HRV) [1].

The HRV is the change in the time intervals between consecutive heartbeats over time, which indicates the heart's adaptation to accept changes in its surrounding [4]. There are various techniques for extracting information from HRV that Poincare plot is one of the techniques commonly used by researchers to qualitatively and quantitatively analyze HRV [5].

The Poincare plot analysis is a geometrical and nonlinear method to assess the dynamics of HRV [6] and is used in sleep studies [7]. It is a diagram in which each RR interval is plotted as a function of the previous RR interval, where the values of each pair of successive RR intervals define a point in the plot [6]. Thus, the plot provides summary information and detailed beat-to-beat information on the behavior of the heart. To optimize the use of the Poincare plot as a quantitative tool, different methods have been suggested.

In this paper, a Poincare plot with ten different lags (plot in which RR_{n+lag} values are plotted against RR_n values) was constructed, and traditional and advanced features were extracted in each lagged plot. Finally, the obtained features were used with a classifier for the detection of sleep apnea.

2. Data and Method

2.1. Data

In this study, two data sets from the Physionet database [8] were used:

1. Apnea- ECG Database [9]
2. Normal Sinus Rhythm Database [8]

The first database contains 70 ECG recordings with apnea annotation [9]. Each record is sampled at 100 Hz and has at least 100 or more minutes with disordered breathing. These recordings were classified into three classes

according to the Apnea-Hypopnea Index (AHI) value. If the recording contained ten or more sleep apnea segments per hour ($AHI \geq 10$) is in class A. If there were five or more sleep apnea segments per hour ($AHI \geq 5$), it belongs to class B. Furthermore, if the recording had less than five segments of sleep apnea per hour ($AHI < 5$) was in class C. In this article, 70 segments were extracted in each class.

To evaluate the extracted features in distinguishing apnea from NSR, 70 ECG segments of subjects in normal sinus rhythm from the second database were also used. The duration of all the recordings was 5 minutes

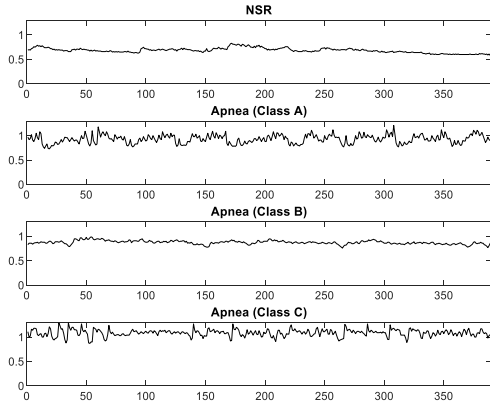


Figure 1. RR intervals of normal sinus rhythm (NSR) and three sleep apnea classes.

2.2. Method

The Pan-Tompkins algorithm was used for QRS detection [10], which was used to find RR intervals (Figure 1). Then extracted features of lagged Poincare plots have been measured for evaluating apnea classes. The features explained in the following subsections were extracted using a Poincare plot of RR intervals in lag equal to one to ten.

2.2.1. Lagged Poincare Plot of RR Intervals

By considering RR time series as $(RR_1, RR_2, \dots, RR_n)$, the lagged Poincare plot is defined by $p_i = (x_i, y_i)$ pairs in which [5]:

$$x_i = (RR_1, RR_2, \dots, RR_{n-lag}) \quad (1)$$

$$y_i = (RR_{1+lag}, RR_{2+lag}, \dots, RR_{n+lag}) \quad (2)$$

where $(i=1, 2, 3, \dots, n)$. Furthermore, n is the number of points in the Poincare plot, and lag represents the value of delay in lagged Poincare plot. Examples of the Poincare plot of sample RR intervals for different apnea classes and NSR for lag equal to one are shown in Figure 2.

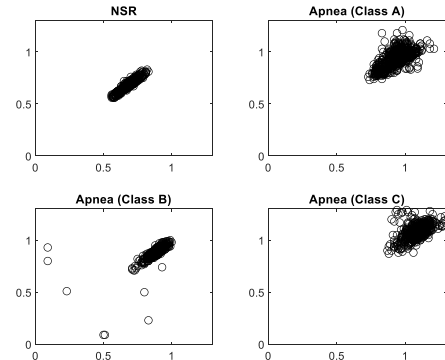


Figure 2. Examples of Poincare plot of RR intervals of NSR and three apnea classes for lag equal to one.

2.2.2. Standard Descriptors of Poincare Plot

$SD1$ and $SD2$ are two standard descriptors of the Poincare plot, which are defined as follows [4]:

$$SD1 = (\text{Var}((x-y) / (2)^{1/2}))^{1/2} \quad (3)$$

$$SD2 = (\text{Var}((x+y) / (2)^{1/2}))^{1/2} \quad (4)$$

Where Var is a variance, the ratio $SD1/SD2$ measures the randomness in the RR intervals time series [11]. Furthermore, $SDNN$ is the standard deviation of RR intervals [11].

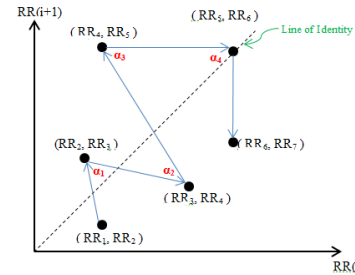


Figure 3. Definition of angle between every three consecutive points in the Poincare plot [11].

2.2.3. New Descriptors of Poincare Plot

For defining the angle between every three consecutive points, points in the Poincare plot were connected in their temporal orders, as shown in Figure 3. The angle between two sequential vectors was measured as explained in [12]. The mean of the obtained angles is used as a feature (mA).

In the Poincare plot, the line passes through the origin at an angle of 45° with a horizontal axis is called the identity line [12]. The distance of points in the Poincare plot relative to the identity line ($y=x$) is determined as follows:

$$d_i = \frac{RR_{i+lag} - RR_i}{\sqrt{2}} \quad (5)$$

The mean of the distance of the points to the identity line is another feature extracted in this research (mD).

2.2.4. Classification

In the classification, 80 percent of the data was used to train a classifier, and 20 percent remained as test data.

For evaluation, the six features extracted from the Poincare plot in lag equal to one were considered first in the classification, then the Poincare plot features in lag 1 to 10 were used for classification, and important features were identified. For evaluation of feature importance, the SHAP (SHapley Additive explanation) framework was used in this study as it is shown to be effective in explaining different classifiers [13].

Table 1. Five-folds cross-validation results for apnea detection using Poincare plot features extracted in the lag equal to one (six features).

	<i>Accuracy</i>	<i>AUC</i>	<i>Recall</i>	<i>Prec.</i>	<i>F1</i>
0	0.756	0.908	0.756	0.770	0.752
1	0.711	0.911	0.708	0.726	0.715
2	0.667	0.869	0.676	0.652	0.653
3	0.578	0.870	0.580	0.614	0.584
4	0.750	0.889	0.750	0.776	0.746
Mean	0.692	0.889	0.694	0.708	0.690
SD	0.066	0.018	0.064	0.065	0.064

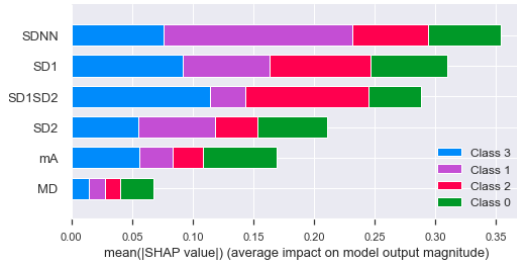


Figure 4. Feature importance plot for apnea classification using features extracted from Poincare plot with lag equal to one. SD1SD2 is the ration of SD1 to SD2, mA is the mean of the obtained angle, and mD is the mean of the distance of the points to the identity line. Green, purple, red, and blue are Class A, Class B, Class C, and NSR, respectively.

2.3. Result

2.3.1. Using features in the lag equal to one

In classifying three apnea classes from NSR using six extracted features of Poincare plot in the lag equal to one, random forest classifier led to the best AUC. AUC for 5-

folds cross-validation is reported in Table 1. Classifier led to an AUC of 0.87 for the test set.

Figure 4 shows feature importance for the classification of NSR versus three sleep apnea classes. As shown, *SDNN* was an important feature in the classification of class B versus other classes. Also, *mA* and *mD* were important in the classification of Class A versus other groups.

2.3.2. Using features in the lags 1-10

As explained, six features have been extracted in each lag of the Poincare plot (60 features). In the classification of three groups of apneas from NSR using 60 extracted features, XGBoost Model [14] led to the best AUC on 5-folds cross-validation, as reported in Table 2. AUC of 0.87 was achieved for the test set.

Figure 5 shows feature importance for the classification of different apnea classes and NSR. Furthermore, selected features in different lags are reported in Table 3. As shown in Figure 5, *SD2 in lag 1*, *SD1 in lag 1*, *SD1/SD2 in lag 1 and 2*, and *mA in lag 1* were the top five important features.

XGBoost selected 25 out of 60 features in the classification task. As shown in Table 3, *mD* and *SD1/SD2* were chosen in all lags to distinguish different sleep apnea classes and NSR.

Table 2. Five-folds cross-validation results for apnea detection using Poincare plot features extracted in the lag equal to one to ten (sixty features).

	<i>Accuracy</i>	<i>AUC</i>	<i>Recall</i>	<i>Prec.</i>	<i>F1</i>
0	0.733	0.894	0.737	0.734	0.731
1	0.756	0.937	0.754	0.768	0.757
2	0.756	0.895	0.759	0.756	0.756
3	0.756	0.914	0.758	0.756	0.754
4	0.705	0.889	0.705	0.722	0.702
Mean	0.741	0.906	0.742	0.747	0.740
SD	0.020	0.018	0.021	0.017	0.021

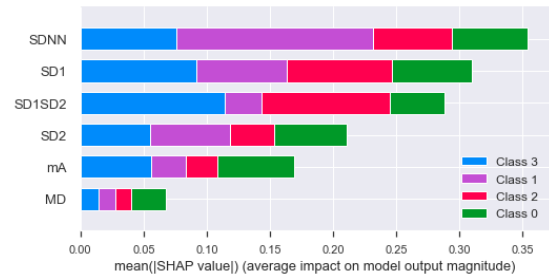


Figure 5. feature importance plot and SHAP values for apnea detection using features of lag-1 Poincare plot

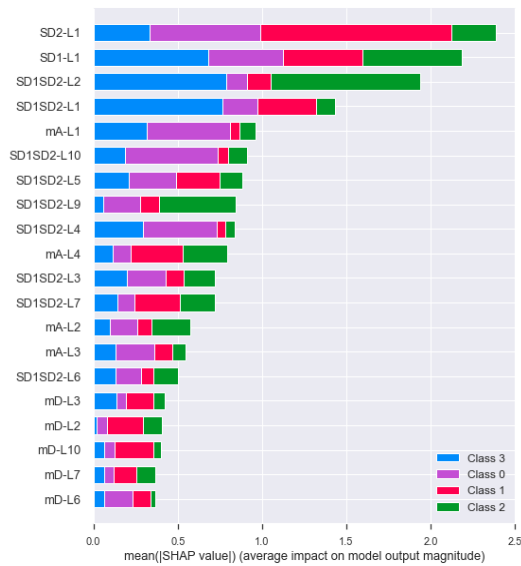


Figure 5. Feature importance plot for apnea classification using features extracted from Poincare plot with lag equal to one to ten. SD1SD2 is the ratio of SD1 to SD2, mA is the mean of the obtained angle, and mD is the mean of the distance of the points to the identity line. Li ($i=1,2,\dots,10$) represents the lag was used for feature extraction. Purple, red, green, and blue are Class A, Class B, Class C, and NSR, respectively.

Table 3. Selected features for classification of sleep apnea classes versus NSR (mA: mean Angle and mD: mean Distance of points to the identity line).

	Lag									
	1	2	3	4	5	6	7	8	9	10
SD1	✓									
SD2	✓									
SD1/SD2	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
SDNN										
mA	✓	✓	✓	✓						
mD	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

2.4. Conclusion

In this research, features extracted from lag 1 to 10 of the Poincare plot have been used for the classification of three classes of apnea from normal sinus rhythm. The results show that Poincare plot-based features in different lags are useful for detecting different apnea classes. Furthermore, by increasing the number of features in different lags, better results are obtained. Furthermore, the new features (angle & distance) have led to promising results for distinguishing apnea in different lags.

References

[1] D.-H. Park *et al.*, "Correlation between the severity of obstructive sleep apnea and heart rate variability indices," *J. Korean Med. Sci.*, vol. 23 (2), pp. 226–231, 2008.

[2] S. Rezaei, S. Moharreri, S. Ghiasi, and S. Parvaneh, "Diagnosis of sleep apnea by evaluating points

distribution in poincare plot of RR intervals," in *Computing in Cardiology*, 2017, vol. 44, doi: 10.22489/CinC.2017.158-398.

[3] V. Cristina, V. Sequeira, P. Bandeira, J. C. Azevedo, and M. Azevedo, *Heart rate variability in adults with obstructive sleep apnea: a systematic review*, vol. 12, 2019.

[4] S. Moharreri, S. Parvaneh, N. J. Dabanloo, and A. M. Nasrabadi, "Utilizing occurrence sequence of heart rate's phase space points in order to discriminate heart arrhythmia," 2010, doi: 10.1109/ICBME.2010.5704919.

[5] M. Brennan, M. Palaniswami, and P. Kamen, "Do Existing measure of poincare plot geometry reflect nonlinear features of heart rate variability?," *IEEE Trans. Biomed. Eng.*, vol. 48 (11), pp. 1342–1347, 2001.

[6] N. Jafarnia Dabanloo, S. Moharreri, S. Parvaneh, and A. M. Nasrabadi, "Application of novel mapping for heart rate phase space and its role in cardiac arrhythmia diagnosis," in *Computing in Cardiology*, 2010, vol. 37.

[7] S. Parvaneh, J. Rubin, A. Samadani, and G. Katuwal, "Automatic detection of arousals during sleep using multiple physiological signals," *Computing in Cardiology*, vol. 45, no. IEEE, pp. 1–4, 2018.

[8] A. L. Goldberger, "Physiobank, physiotoolkit, and physionet components of a new researchresource for complex physiologic signals," *Circulation*, vol. 101, no. 2, pp. e215–e220, 2000.

[9] T. Penzel, G. B. Moody, R. G. Mark, A. L. Goldberger, and J. Peter, "The Apnea-ecg database," *Comput. Cardiology*, 2000.

[10] J. Pan and W. J. Tompkins, "A real time qrs detection algorithm," *IEEE Trans. Biomed. Eng.*, vol. 32 (3), pp. 230–236, 1985.

[11] F. Shaffer and J. P. Ginsberg, "An overview of heart rate variability metrics and norms," *Front. Public Heal.*, vol. 5, 2017.

[12] S. Moharreri, N. J. Dabanloo, S. Rezaei, and S. Parvaneh, "New feature set for better representation of dynamic of RR intervals in poincare plot," in *Computing in Cardiology*, 2017, vol. 44, doi: 10.22489/CinC.2017.321-395.

[13] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *Adv. Neural Inf. Process. Syst.*, pp. 4765–4774, 2017.

[14] Chen *et al.*, "XGBoost: extreme gradient boosting," *R Packag. Version 0.4-2 1*, vol. 4, pp. 1–4, 2015.

Address for correspondence.

Saman Parvaneh

1 Edwards Way, Irvine, CA, USA

parvaneh@ieee.org