ECG Signal Analysis on an Embedded Device for Sleep Apnea Detection

Rishab Khincha(✉) , Soundarya Krishnan , Rizwan Parveen ,
and Neena Goveas

Department of Computer Science, BITS Pilani Goa Campus, Sancoale 403726, India
khincharishab@gmail.com, soundaryak4898@gmail.com,
{rizwanp,neena}@goa.bits-pilani.ac.in

Abstract. Low cost embedded devices with computational power have the potential to revolutionise detection and management of many diseases. This is especially true in the case of conditions like sleep apnea, which require continuous long term monitoring. In this paper, we give details of a portable, cost-effective and customisable Electrocardiograph(ECG) Signal analyser for real time sleep apnea detection. We have developed a data analysis pipeline using which we can identify sleep apnea using a single lead ECG signal. Our method combines steps including dataset extraction, segmentation, signal cleaning, filtration and finally apnea detection using Support Vector Machines (SVM). We analysed our proposed implementation through a complete run on the MIT-Physionet dataset. Due to the low computational complexity of our proposed method, we find that it is well suited for deployment on embedded devices such as the Raspberry Pi.

Keywords: ECG · Embedded device · Signal processing · Machine learning · Sleep apnea

1 Introduction

Sleep apnea is a condition that occurs when the upper airway becomes blocked during sleep, reducing or completely stopping airflow, or if the brain fails to send the signals needed to breathe. This is a condition that could prove life-threatening if not monitored continuously and automatically.

Sleep apnea diagnosis and treatment is difficult, as the diagnosis involves overnight polysomnography (PSG), during which a medical expert is required to work overnight [1]. Thus, for medical conditions like sleep apnea, setting up automated monitoring and detection of anomalies is a crucial requirement. Medical practitioners use expensive hospital based equipment and in-patient monitoring for detection of anomalies. In addition to being expensive, this kind

of monitoring in an artificial setting for conditions like sleep apnea may not give accurate results. The cost and the discomfort may result in many patients avoiding treatment until the condition becomes advanced and life threatening. Furthermore, these devices are not customisable by users in order to add more features or modify existing ones.

Electrocardiography (ECG) signals are one of the most feature rich and non-intrusive ways to detect various cardiac disorders [2]. ECG sensors and cost effective and powerful embedded devices are now freely available. The only thing lacking are mechanisms which can lead to real time analysis of data captured using these devices. Wang et. al. presented an AI based mechanism to detect sleep apnea from a single line ECG [3]. In this work, we propose a simpler algorithm to detect sleep apnea, that uses classifiers to enable execution with required speed on a device such as a Raspberry Pi. With its low cost and portability, any user can buy a device and monitor sleep apnea in the natural setting of their homes. In addition, our device will also be customisable for the users to add additional software to make the most of the device. In this work, we present a data analysis pipeline for raw ECG signals to be used for real time on-device detection of this condition. We find that our data analysis pipeline works with good accuracy with the required real-time efficiency. Our device can have utility for many other use cases requiring continuous monitoring and detection ranging from driver-fatigue detection and health anomaly detection in soldiers deployed in inhospitable locations.

In this work, we study the problem of development of a data analysis pipeline which can be executed on a resource constrained embedded device. In Sect. 3, we describe the steps that are part of a complete end to end pipeline starting from raw signals to an alert based actuation mechanism.

2   ECG Signals

2.1   ECG Signal Features

ECG signals consist of five types of waves, i.e., the P, Q, R, S and T waves. The parts of a signal are shown in Fig. 1, and the medical relevance of these parts is summarised in Table 1.

Table 1. Features of ECG signals

|	Time from	Time to	What it indicates	Conditions associated
RR interval	R peak	Next R peak	ECG signal period	Sleep apnea
P-R interval	Start of P wave	Start of QRS complex	Impulse time from sinus node to ventricles	Heart blockage
Q-T interval	Start of Q wave	End of T wave	Ventricular (De/Re)polarization	Cardiac death
QRS complex	Q, R, S waves		Ventricular repolarization	Electrolyte imbalance, Drug toxicity
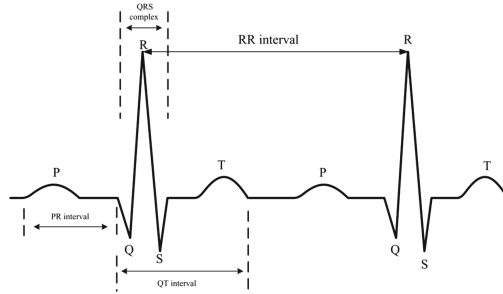
Fig. 1. Parts of an ECG signal

Irregular heartbeats lead to abnormal waves which can be traced from the patient's ECG signals. These irregular rhythms of the heart can be used to detect sleep apnea and other important medical conditions [4]. Rhythmic cycles during a heartbeat, especially the RR intervals in ECG signals have been reported to be associated with sleep apnea.

ECG signals can now be easily measured using commercially available embedded devices [5]. Long term monitoring of these signals is now possible using newly designed innovative power efficient devices, sensors and electrodes which enable ease of use [6]. These efforts have led to the possibility of designing cost effective and customised ECG signal capturing and analysis devices. The need now is to automate the analysis of the collected signals to detect various anomalies.

2.2   ECG Signals Dataset

MIT-Physionet database is the standard database used by researchers around the world for studies involving ECG signals [7]. We have used the Apnea-ECG Database for our work. [8]. The dataset contains 70 records, out of which only 35 have apnea annotations. Recordings vary in length from slightly less than 7 h to nearly 10 h. Each recording includes a continuous digitised ECG signal, a set of apnea annotations (derived by human experts), and a set of machine-generated QRS annotations. The digitised ECG signals (16 bits per sample) are recorded at the rate of 100 samples per second. A subset of the dataset was used for training. 17125 min of ECG recordings, at a rate of 100 Hz was used for training. Out of this, about 6514 min are apneatic, which gives us a well balanced training set. The training set has a 6:1 male to female ratio.

3   Data Pipeline

Raw signals as obtained from any device have several imperfections which need to be dealt with before being used to detect any anomalies. The digitised ECG signals have been recorded at the rate of 100 samples per second. These signals are then processed through the pipeline shown in Fig. 2.
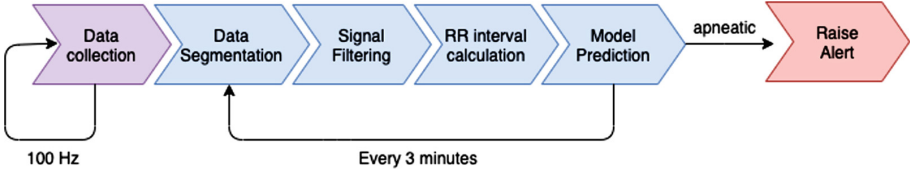
Fig. 2. Pipeline for our device

The data is first segmented into blocks of customisable duration in order to generate features. The signal is then filtered and cleaned as shown in Sect. 3.2 to remove respiratory artefacts and allow for suppression of non-R peaks. The RR intervals are calculated from the filtered signal and various data features are generated. Finally, the SVM model is used to predict whether the patient was apneatic in the last 3 min or not, and an alert is raised if true.

3.1 Data Segmentation

The dataset is segmented into blocks comprising of a few minutes each to perform analysis. Further, we discard the first few and the last few samples in every block to remove edge irregularities, if any. Each block can be customised to accommodate samples for a few minutes starting from one minute long blocks, since the MIT-Physionet dataset has annotations for every minute. A smaller block size gives more frequent predictions while a larger block size gives better statistical insights. In this work, we have used a block size of 3 min which gives us an average of approximately 200 RR intervals for feature generation. In our pipeline, the block duration is customisable to suit the needs of the user.

3.2 Signal Cleaning and Filtering

ECG Signal capturing devices like Holter Monitors continuously collect data over large periods of time. Part of this data acquired over long periods of time may be corrupted due to patient movement, sensor placement, and interference from other sources. As discussed by Mardell et al. [9], algorithms that are computationally less intensive are heavily dependant on the quality of signals provided. Drift artefacts caused due to breathing usually lie around 0.5 Hz while motion artefacts caused due to human motion lie around the 5 Hz range. Filtering using a high pass filter of any frequency above 5 Hz removes these artefacts. Since our analysis uses RR intervals, we found that a high pass filter of 20 Hz gave the best suppression of non-R peak segments of the ECG signal.

Figure 3 is a sample three minute segment of the ECG data without filtering which shows some of the artefacts mentioned above. This signal is then filtered through a high pass filter of 20 Hz causing the artefacts to be removed as shown in Fig. 4.

RR intervals are defined as $RR(i) = R(i+1) - R(i), i = 1, 2, ..., n-1$ where $R(i)$ is the time at which the ith R peak occurs. Thireau et al. [10] have shown
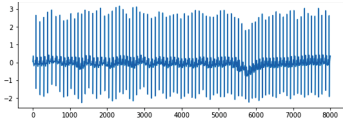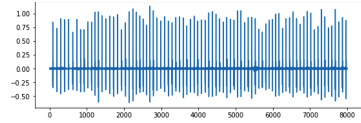
Fig. 3. ECG segment

Fig. 4. ECG segment after filtering

with a 95% confidence interval that the R peaks are at least two times the standard deviation above the mean value of the signal. Our pipeline uses this result for R peak detection, thus using minimal computational power and running time.

When calculating the RR intervals, it is possible that the algorithm discussed fails to detect some of the R peaks, resulting in some abnormally long RR intervals. [11]. To avoid this, we remove the abnormal RR intervals by replacing them with the mean RR interval. The final result of filtering before the intervals are extracted helps us easily single out the R peaks by suppressing all the other peaks, as can be seen in Fig. 6. The features generated from the RR intervals, as proposed by Chazal et al. [12] and Isa et al. [13] are seen to be most effective in detecting conditions of sleep apnea. Our model uses the following features: mean, standard deviation and median of RR intervals, NN50 measures (number of pairs of successive RR intervals that differ by more than 50 ms), and the inter quartile range of the RR interval distribution (Fig. 5).
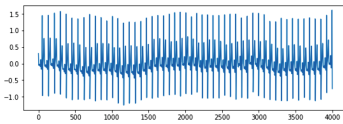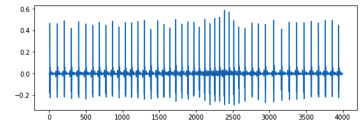

Fig. 5. ECG segment

Fig. 6. ECG segment after suppression

3.3 Apnea Detection

Figure 7 shows a sample segment of a subject showing normal breathing pattern, while Fig. 8 shows a similarly sized segment showing sleep apnea. When a person is unable to breathe, it increases the body's 'fight-or-flight' stress response, making the heart beat faster, and thus, makes the RR intervals shorter.

We first attempt a comparative study of some algorithms that require computational power lesser than the one provided by an embedded device like the Raspberry Pi to determine the optimal one for the problem at hand. The comparative results are shown in Table 2.

As seen from Table 2, Support Vector Machines gave an accuracy of 87.23% with a good precision and recall. This algorithm, with an F1 score of 0.897 works well in identifying sleep apnea. Figure 9 depicts the confusion matrix, i.e.
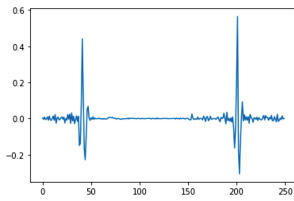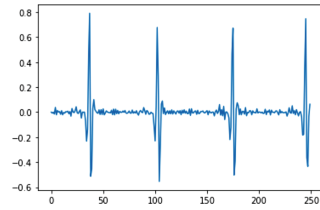
Fig. 7. Non-apneatic segment

Fig. 8. Apneatic segment

Table 2. Comparison of different algorithms

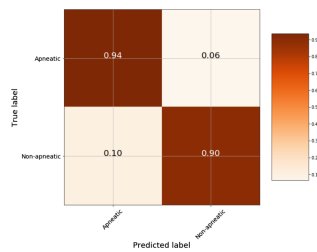| Algorithm	Accuracy	Precision	Recall
SVM	87.23%	89.34%	90.14%
Random forest	81.80%	80.39%	86.32%
KNN	84.11%	86.14%	85.62%
Decision tree	85.62%	83.56%	82.18%

Fig. 9. Confusion Matrix for SVM with a Radial Basis Function (RBF) kernel

the True label vs. the Predicted label for our trained SVM model. We find that the probability of a false negative, where an apneatic episode is failed to be detected is 0.06.

3.4   Timings

Considering the scenario in which blocks of ECG signal data are being appended into a file constantly by the capturing device, our code reads the data written in every block duration and feeds it to our pipeline. The pipeline processes these signals and makes a prediction about the presence or absence of sleep apnea, which is then written onto a file.

We ran 584 h of ECG data from the MIT-Physionet dataset through this pipeline on a Raspberry Pi 3 Model B and a 2.5 GHz Intel Core i7-6500U with 8 GB RAM laptop. Table 3 shows the time statistics for running all the blocks through the pipeline, including the reading and the writing times.

Table 3. Comparison of running time on a laptop vs. a Raspberry Pi

	Signal length (hours)	Mean runtime (s)	Standard deviation (s)
Laptop	584	0.229	0.026
Raspberry Pi	584	2.43	0.0953

3.5   Results

We find that our proposed pipeline can perform end to end analysis of ECG data within 2.5 s on a Raspberry Pi for a 3 min segment when simulated for a 584 h dataset. This gives us an idle time of about 177.5 s every 180 s. This ensures that all the data does not need to be transmitted for analysis. Our proposed pipeline can identify apneatic episodes and only necessary information about the sleep apnea episodes needs to be sent. In addition, the apneatic episode related data can be transmitted within a few seconds using cellular transmission or a Wi-Fi based sensing network as shown by Yang et al. [14].

A comprehensive solution to the identification and treatment of sleep apnea includes actuation of devices attached to the individual. This will ideally result in halting of the episode. Some of the previously proposed mechanisms include vibration of a device [15] and physical stimulation of muscles [16]. We find that using our pipeline executing on a Raspberry pi, there is idle processor time available which can be used to raise alerts and actuate devices as required.

4   Conclusion

In this paper we have introduced a software pipeline suitable for real-time sleep apnea detection using ECG signals which can run on a resource constrained embedded device. Our method combines data segmentation, signal filtering and feature generation from RR intervals to predict occurrence of sleep apnea. We find that our pipeline successfully predicts conditions of sleep apnea on a Raspberry Pi with an execution time of only two percent of the time duration of the data. Our proposed system is portable, customisable and would be an ideal solution for a continuous monitoring device for sleep apnea. In the future, we plan to include a more detailed analysis of the ECG signals, actuation mechanisms, comparison with the user's historical data and incorporation of additional sensors for measuring other body parameters. This opens up the possibility of a portable, cost-effective, customisable health monitoring device.

References

1. Almazay L., Elleithy, K., Faezipour, M.: Obstructive sleep apnea detection using SVM-based classification of ECG signal features. In: 34th Annual International Conference of the IEEE EMBS San Diego, California USA, 28 August–1 September (2012)

2. Benazza-Benyahia, A., Ben Jeb, S.: Multiresolution based reference estimation for adaptive ECG signals denoising. In: International Conference on Image and Signal Processing ICISP, Morocco, vol. 2, pp. 875–882 (2001)

3. Wang, X., Cheng, M., Wang, Y., et al.: Obstructive sleep apnea detection using ECG-sensor with convolutional neural networks. Multimed. Tools Appl. (2018). https://doi.org/10.1007/s11042-018-6161-8

4. Huang, R., Zhou, Y.: Disease classification and biomarker discovery using ECG data. https://doi.org/10.1155/2015/680381

5. Sreekesh, S., Abhimanyu, Z., Keerthi, G., Goveas, N.: Customizable Holter monitor using off-the-shelf components. In: ICACCI, pp. 2302–2306. IEEE Explore (2016). https://ieeexplore.ieee.org/document/7732396

6. Braojos, R., et al.: Ultra-low power design of wearable cardiac monitoring systems. In: Conference Proceedings of Annual Design Automation Conference, vol. 1, pp. 1–6 (2014)

7. Goldberger, A.L., et al.: PhysioBank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. Circulation **101**(23), e215–e220 (2003)

8. Penzel, T., Moody, G.B., Mark, R.G., Goldberger, A.L., Peter, J.H.: The Apnea-ECG database. In: Computers in Cardiology, vol. 27, pp. 255–258 (2000)

9. Imtiaz, S.A., Mardell, J., Saremi-Yarahmadi, S., Rodriguez-Villegas, E.: ECG artefact identification and removal in mHealth systems for continuous patient monitoring. Healthc. Technol. Lett. **3**(3), 171–176 (2016). https://doi.org/10.1049/htl.2016.0020

10. Thireau, J., Zhang, B.L., Poisson, D., Babuty, D.: Heart rate variability in mice: a theoretical and practical guide. Exp. Physiol. **93**(1), 83–94 (2008)

11. Karlsson, et al.: Automatic filtering of outliers in RR intervals before analysis of heart rate variability in Holter recordings: a comparison with carefully edited data. BioMed. Eng. (2012). https://doi.org/10.1186/1475-925X-11-2

12. Chazal, P., Penzel, T., Heneghan, C.: Automated detection of obstructive sleep apnea at different time scales using the electrocardiogram. Inst. Phys. Publ. **25**(4), 967–983 (2004)

13. Isa, M., Fanany, S., Jatmiko, M.I., Aniati, W.A.: Feature and model selection on automatic sleep apnea detection using ECG, pp. 357–362 (2010). https://doi.org/10.13140/RG.2.1.1624.7124

14. Yang, Z., Zhou, Q., Lei, L., Zheng, K., Xiang, W.: An IoT-cloud based wearable ECG monitoring system for smart healthcare. J. Med. Syst. **40**(12), 1–11 (2016). https://doi.org/10.1007/s10916-016-0644-9

15. Scarlata, S., Rossi Bartoli, I., Santangelo, S., Giannunzio, G., Pedone, C., Antonelli Incalzi, R.: Short-term effects of a vibrotactile neck-based treatment device for positional obstructive sleep apnea: preliminary data on tolerability and efficacy. J. Thorac. Dis. **8**(7), 1820–1824 (2016). https://doi.org/10.21037/jtd.2016.04.69

16. Spicuzza, L., Caruso, D., Di Maria, G.: Obstructive sleep apnoea syndrome and its management. Therap. Adv. Chronic Dis. **6**(5), 273–285 (2015). https://doi.org/10.1177/2040622315590318