Sleep Apnea Detection by a Recurrent Neural Network based on Long Short-Term Memory

Keita Nishio\textsuperscript{a,}\textsuperscript{*}, Takashi Kaburagi\textsuperscript{b}, Satoshi Kumagai\textsuperscript{a}, Toshiyuki Matsumoto\textsuperscript{a} and Yosuke Kurihara\textsuperscript{a}

\textsuperscript{a}Aoyama Gakuin University, 5-10-1 Fuchinobe, Chuo-ku Sagamihara-city, Kanagawa, 252-5258, Japan
\textsuperscript{b}International Christian University, 3-10-2 Osawa, Mitaka-city, Tokyo, 181-8585, Japan

\*Keita Nishio: nishio.engineering.agu@gmail.com

Abstract

Sleep apnea syndrome is one of the most serious sleep disorders. Sleep apnea detection systems have been proposed which estimate whether the measured signal contains an apnea event. Hence, we proposed apnea state detection at each discrete time by utilizing frequency distribution. The objective of this study was to improve the detection accuracy through the consideration of the temporal association of the frequency distribution.

A pneumatic bio-instrumental system was adopted as the sensing system. The output signal contained biomedical information such as respiration, heartbeat, and body movement. Wavelet transform was applied to the output signal to obtain the temporal alteration of the frequency distribution. Wavelet coefficients in the range of 0.15Hz to 0.45Hz were used as the respiratory distribution, which was used as an input in the recurrent neural network. A long short-term memory layer in the recurrent neural network utilized the last state as an update which enabled the network to take into account the temporal association. The output from the long short-term memory layer was interpreted using the probabilities that passed through a fully connected layer and a softmax layer. The apnea state at each discrete time was detected by comparing the probabilities.

To validate the method, an experiment in which 8 subjects participated was conducted. The results indicate the detection performance of the proposed method is better than the previous method in all three indices. It is inferred that the temporal alteration of respiratory distribution is effective for apnea detection.

**Keywords:** sleep apnea, frequency distribution, long short-term memory.

1. Introduction

According to the World Health Organization, ischemic heart disease (IHD) was the leading cause of death by disease in 2016\textsuperscript{(1)}. Recent investigations reported a relationship between IHD and sleep apnea syndrome (SAS). Marin, Carrizo, Vicente, and Agusti (2005) reported that untreated sleep apnea-hypopnea significantly increased the risk of a fatal cardiovascular event caused by IHD\textsuperscript{(2)}. Moreover, Maia, Goulart, Drager, Staniak, Santos, Lotufo, and Bensenor (2017) reported that 30\% of the IHD patients are comorbid with SAS\textsuperscript{(3)}. However, the diagnosis rate of SAS is estimated to be only 15\% because it is very difficult for the SAS patients to recognize that they have SAS.

Taking these in consideration, sleep apnea detection systems have been proposed\textsuperscript{(4-9)}\textsuperscript{). However, these systems focus on whether the measured signal contains an apnea event. To evaluate the number of apnea events, we propose an apnea detection system\textsuperscript{(10)} at each discrete time in which a frequency distribution is utilized to detect the occurrence of an apnea event. In this paper, in order to improve detection accuracy, we propose a sleep apnea detection which takes into account the temporal association of the frequency distribution.

2. Proposed Method

This chapter describes a sleep apnea detection method based on a recurrent neural network (RNN) that utilizes the temporal association of a frequency distribution in the output signal from a pneumatic sensing system\textsuperscript{(11)}. The entire signal processing consists of a learning phase and a detection phase.
2.1 Learning Phase

Let \( k=1,2,3,\ldots,K \), \( x(k) \) and \( l(k) \in (0,1) \) be the discrete time step, the output signal from the sensing system, and a state where 1 and 0 represent a respiration state and an apnea state at each time, respectively. \( K \) is the number of the output signal. All the output signals \( x(k) \) obtained in the learning phase are labeled by \( l(k) \).

The sensing system is placed under an ordinary bed mattress, measures a pressure change, and produces an electric signal \( x(k) \) as output. The output signal \( x(k) \) contains biosignals such as respiration, heartbeat and body movement.

Wavelet transform (WT) is applied to \( x(k) \) to obtain the temporal alteration of the respiratory distribution \( R(k,f) \), where \( f=1,2,3,\ldots,F \) is the discrete frequency. \( F \) represents the number of frequency data points for each time step \( k \). Assuming an ordinary respiratory frequency between 0.15Hz and 0.45Hz, the frequency range of the WT is set to the respiratory frequency band. The respiratory distribution \( R(k,f) \) represents the temporal alteration of the frequency distribution in the respiratory frequency band.

To utilize the temporal alteration in apnea detection, an RNN based on Long Short-Term Memory (LSTM) is constructed. The whole network consists of an input layer, an LSTM layer, a fully connected layer, and a softmax layer in sequence. The dimensions of the input layer and the LSTM layer are \( D_{i} \) and \( D_{o} \). The dimension of the input layer \( D_{i} \) is equal to \( F \). The state of the LSTM layer \( s(k) \) is updated by both the respiratory distribution \( R(k,f) \) and the previous state \( s(k-1) \). Thus, the LSTM layer takes into account the temporal alteration of the respiratory distribution \( R(k,f) \). The output from the LSTM layer is the input of the fully connected layer where a two-dimensional output set corresponds to a respiration state and apnea state. The softmax layer converts the two-dimensional output to two probabilities: the probability of the respiration state \( p_{r}(k) \) and the probability of the apnea state \( p_{a}(k) \).

In the learning phase, the network is trained to produce high \( p_{r}(k) \) for data labeled \( l(k)=1 \). Similarly, the network is also trained to produce high \( p_{a}(k) \) for data labeled \( l(k)=0 \). The learning process is executed based on the RMSProp optimization algorithm.

2.2 Detection Phase

In the detection phase, an apnea state in the measured signal \( x(k) \) is detected at each discrete time. Let \( x_{d}(k) \) and \( R_{d}(k,f) \) be the measured signal and the \( x(k) \) and the respiratory distribution \( R(k,f) \) in the detection phase, respectively. The label \( l_{d}(k) \) of \( x_{d}(k) \) is unknown. The respiratory distribution \( R_{d}(k,f) \) is inputted to the trained neural network. The neural network gives the probability of respiration state \( p_{r}(k) \) and the probability of apnea state \( p_{a}(k) \). If \( p_{r}(k) > p_{a}(k) \), \( l_{d}(k) \) is determined to be 0 (apnea state). Otherwise, \( l_{d}(k) \) is determined to be 1 (respiration state).

3. Experiment

This chapter describes the experiment conducted to evaluate the proposed method.

3.1 Experimental Environment

Fig. 1 shows a whole environment for the experiment.

The sensing device was placed under an ordinary mattress 60 cm from the head of the bed so that the device was located directly under the heart of a subject. Multifunction I/O Device USB-6000 (National Instruments Corporation) was used for A/D conversion. The sampling interval was 10ms and the measurement time was 180s.

3.2 Experimental Procedure

All experiments were performed in accordance with the program checked and approved by the Life Science Committee of the Aoyama Gakuin University (permission No. M15-17). Moreover, informed consent was obtained before the initiation of the experiment. Each subject was asked to do four possible sleeping positions: a supine position, a prone position, a right lateral position, and a left-lateral position. One data item was obtained for each sleeping position, so four data items were obtained from each subject. Eight subjects participated in this experiment, so the total number of date items was 32.

The subject lay on the mattress in the designated position. Before beginning measurement, a 40-second rest period was set because the respiration was unsteady immediately after lying on the mattress. After the rest period,
data measurement for 180 seconds began. Each subject stopped breathing three times for 20 seconds: 20s to 40s, 80s to 100s, and 140s to 160s from the beginning of the measurement period. The state during the arrest period was labeled \( l(k)=0 \). Moreover, the mother wavelet was Morlet, \( F=16 \) and the dimensions of the input layer and the hidden layer were \( D_i=16 \) and \( D_h=50 \).

### 3.3 Evaluation Criteria

We utilized two cross validation techniques: leave-one-subject-out cross validation and leave-one-position-out cross validation. The same cross validation tools were used in our previous method\(^{(10)}\) as well. Comparing the results of the proposed method and the previous method, we examined the effectiveness of considering the temporal association.

In the leave-one-subject-out cross validation, one subject was selected as test subject while the remaining subjects were used as training subjects. For the data of the training subjects, the respiratory distribution \( R(k,f) \) was calculated and the whole neural network was trained with the correct label \( l(k) \). The respiratory distribution \( R(k,f) \) was likewise calculated using the data of the test subject. The respiratory distribution \( R(k,f) \) was inputted to the trained neural network to detect apnea state. Finally, we classified the detection results as the true positive (\( TP \)), true negative (\( TN \)), false positive (\( FP \)), and false negative (\( FN \)). Moreover, the following three evaluation indices were calculated based on Eq. (1) to (3) for the test subject: sensitivity (\( Sens \)), specificity (\( Spec \)), and accuracy (\( Acc \)).

\[
\text{Sens} = \frac{TP}{TP + FN} \quad (1)
\]
\[
\text{Spec} = \frac{TN}{TN + FP} \quad (2)
\]
\[
\text{Acc} = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)
\]

For this evaluation criteria, leave-one-subject-out cross validation was conducted for all combinations of the test subject and training subjects. Therefore, the above three evaluation indices were calculated for each test subject.

In the leave-one-position-out cross validation, one sleeping position was selected as the test position and the remaining sleeping positions were used as training positions. The training of the RNN and the evaluation indices follow the leave-one-subject-out cross validation.

### 4. Result

This chapter describes the result of the validation experiment. Section 4.1 shows examples of signal processing. Section 4.2 shows the evaluated results of the two cross validations.

#### 4.1 Example of Signal Processing

The \( x(k) \) of ID7 in each sleeping position is shown in Fig. 2 to Fig. 5 Each figure shows (a) the output signal \( x(k) \), and (b) the respiratory distribution \( R(k,f) \).

With regard to the output signal \( x(k) \), clear characteristics of respiratory arrest cannot be seen in any of the sleeping positions. On the other hand, the spectrum of the respiration distribution \( R(k,f) \) appears to decrease during the arrest periods especially in a supine position and a left lateral position. However, those characteristics remain obscure.

![Fig. 2. \( x(k) \) and \( R(k,f) \) of supine position.](image)

![Fig. 3. \( x(k) \) and \( R(k,f) \) of prone position.](image)
4.2 Evaluation Results

Fig. 6, Fig. 7 and Fig. 8 show Sens, Spec and Acc by the leave-one-subject-out cross validations. Each figure shows the evaluation result for each subject from (a) Proposed method and (b) Previous method. The average of Sens, Spec, and Acc from the previous method was 0.82, 0.88 and 0.86. On the other hand, the proposed method achieved an average of 0.86, 0.89 and 0.88, respectively. The proposed method outperformed the previous method in all the indices.

Moreover, Fig. 9 to Fig. 11 show Sens, Spec and Acc by leave-one-position-out cross validation. The average of Sens, Spec, and Acc by the previous method are 0.82, 0.90 and 0.87. Meanwhile, the proposed method achieved an average of 0.87, 0.84 and 0.92, respectively. The proposed method outperformed the previous method in two indices: Sens and Acc.
state does not clearly appear; so, the previous method cannot detect apnea state with high accuracy. In the proposed method, because the temporal association is taken into account, it is possible to detect apnea state even if the apnea state does not clearly appear as shown in Fig. 3 and Fig. 4.

6. Conclusion

In this paper, we developed sleep apnea detection by an RNN based on LSTM which focuses on the temporal alteration of the respiratory distribution. The validation experiment yields that the proposed method outperformed the previous method. The results indicate the temporal alteration of respiratory distribution is effective for sleep apnea detection. This research focused on respiratory information in a low-frequency domain. Further studies are needed to utilize the respiratory information in a sound domain as well.

References


(6) C. Yang, G. Cheung, V. Stankovic, K. Chan and N. Ono: “Sleep Apnea Detection via Depth Video and Audio


