

RESEARCH ARTICLE

# Integrating medical rules to assist attention for sleep apnea detection

Jianqiang Li<sup>1</sup>, Xiaoxiao Song<sup>1</sup> , Yanning Lin<sup>1</sup>, Junya Wang<sup>1</sup>, Dongying Guo<sup>2</sup> and Jie Chen<sup>1</sup> 

<sup>1</sup>College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China and <sup>2</sup>Department of Respiratory and Critical Care Medicine, Shenzhen People's Hospital, Shenzhen, China

**Corresponding author:** Jie Chen; Email: [chenjie@szu.edu.cn](mailto:chenjie@szu.edu.cn)

**Received:** 30 June 2022; **Revised:** 27 February 2023; **Accepted:** 24 March 2023; **First published online:** 28 April 2023

**Keywords:** sleep apnea detection; self-attention; medical rule; interpretable models

## Abstract

Sleep apnea is one of the most common sleep disorders. The consequences of undiagnosed sleep apnea can be very serious, increasing the risk of high blood pressure, heart disease, stroke, and Alzheimer's disease over a long period of time. However, many people are often unaware of their condition. The gold standard for diagnosing sleep apnea is nighttime polysomnography monitoring in a specialized sleep laboratory. However, these diagnoses are expensive and the number of beds is limited, and there is insufficient monitoring in terms of time dimension. Existing methods for automated detection use no more than three physiological signals, but all other signals are also associated with the patient's sleep. In addition, the limited amount of medical real annotation data, especially abnormal samples, lead to weak model generalization capability. The gap between model generalization capability and medical field needs still exists. In this paper, we propose a method for integrating medical interpretation rules into a long short-term memory neural network based on self-attention with multichannel respiratory signals as input. We obtain attention weights through a token-level attention mechanism and then extract key rules of medical interpretation to assist the weights, improving model generalization and reducing the dependence on data volume. Compared with the best prediction performance of existing methods, the average improvements of our method in accuracy, precision, and f1-score are 3.26%, 7.03%, and 1.78%, respectively. The algorithm tested the performance of our model on the Sleep Heart Health Study data set and found that the model outperformed existing methods and could help physicians make decisions in their practices.

## 1. Introduction

Sleep apnea (SA) is a sleep-related disease, and it is characterized by difficulty in breathing during sleep [1, 2]. The disease can be divided into two categories by its etiology: (1) obstructive sleep apnea (OSA) that is caused by obstruction of the airway by the throat muscles [3] and (2) central sleep apnea (CSA) which is caused by a disturbance in the brain center that controls breathing [4]. People of all ages are at risk of SA. Approximately 200 million people (4% of adult men and 2% of adult women) [5] in the world suffer from sleep-disordered breathing [6, 7]. According to report [8, 9], in the United States, 93% of middle-aged women with SA and 82% of patients with moderate to severe SA are undiagnosed. Studies [10] have also shown that the prevalence rate of preschool children is 3%. Moreover, SA is associated with ischemic heart disease, cardiovascular dysfunction and stroke [11], daytime sleepiness [12], and could be related to the development of diabetes mellitus type 2 (T2DM) [13].

Currently [14], the gold standard for diagnosing sleep apnea is all-night polysomnography (PSG) in the sleep laboratory [1]. To enable doctors to obtain accurate results [15], PSG records involve at least 11 channels of various physiological signals collected from different sensors, including electroencephalogram (EEG), electrooculogram (EOG), electromyography (EMG), and electrocardiogram (ECG), etc [16]. Due to a large number of sensors mounted to the body, the patients tend to feel uncomfortable

[17]. In addition, the PSG service is normally expensive and unavailable for most people [12]. The analysis process is time-consuming and laborious [18]. Generally, the qualified professionals who can diagnose sleep apnea in medical institutions are very limited [13]. Therefore, there is an urgent need to automatic SA detection [19] and help technicians achieve high accuracy and throughput in SA diagnosis [20].

Deep learning has a wide range of applications in the medical field [21, 22]. For example, Zhou et al. [23, 24] demonstrated a robust framework for needle detection and localization in subretinal injection using microscope-integrated optical coherence tomography based on deep learning. Park et al. [25] proposed a frequency-aware based attention-based LSTM (long short-term memory) for cardiovascular disease that weighs on important medical features using an attention mechanism that considers the frequency of each medical feature. Various automatic methods have been proposed to help diagnose SA. Steenkiste et al. [26] proposed an automatic SA detection method based on LSTM neural networks, which uses the original physiological respiratory signals to automatically learn and extract related characteristics, and to detect possible sleep apnea events. The authors use balanced bootstrapping for the experiments to be conducted each time using an entire minority class and majority classes of the same size. The method achieved an average true positive rate of 80% by using three sensor signals, including abdominal respiratory, thoracic respiration and ECG-derived respiration (EDR). Thorey et al. [27] proposed a fully convolutional and highly parallelizable method based Convolutional Neural Network 1D (CNN1D) that can process signals of any sizes efficiently. Their method reached an average accuracy 81% for sleep apnea severity diagnosis by using more physiological signals. However, existing researches suffer from three limitations: (1) PSG involves multiple signals, but most of the existing methods are based on no more than three signals while all other signals are fully utilized. (2) The amount of labeled data are limited, especially in abnormal samples, which leads to poor generalization ability. (3) The accuracy of the current algorithm still needs to be improved for practical usage.

To address the above limitations, this work proposes a method which can integrate domain knowledge in the form of medical rules into LSTM neural network which can utilize multichannel respiratory signals based on self-attention mechanism. In this work, we obtain the attention weight through the word-level attention mechanism and then extract the key medical rules from the doctors and place them on the input to obtain the auxiliary weights. Subsequently, the proposed method connects the two weights through a real-valued hyperparameter to guide the attention values. Finally, the hyperparameter is optimized by Bayesian optimization (BO) to obtain a model with better generalization capability.

Toward development of automatic SA detection, the contributions of this work can be summarized as follows:

- The proposed method can detect SA by using all signals (including ECG, EEG, thoracic respiratory, etc.) in PSG as multichannel inputs to model data (Section 3.2). The results demonstrate that the effect of multichannel input is superior to that of conventional three-channel input and any single-channel input.
- The proposed method integrates the medical rules into model to assist the attention weight, which can improve model generalization and effectively alleviate the dependence on the amount of data in the case of reduced data volume (Section 3.3).
- The proposed method is tested on the publicly available Sleep Heart Health Study dataset and it is shown that our model outperforms existing methods and can help physicians make decisions in practice (Section 4.4.2).

## 2. Related work

### 2.1. Automatic sleep apnea detection

Previous works have tried to automatically detect sleep apnea using deep neural network (DNN) models, such as LSTM neural networks and convolution neural networks (CNN). Steenkiste et al. [26] used an

LSTM neural network to capture temporal information and accurately model the data. A fourth-order low-pass zero-phase shift Butterworth filter was first used to reduce noise in the respiratory signal and automatically predict OSA events based on the expansion and contraction patterns of abdominal respiration, thoracic respiration, and EDR. Haidar et al. [28] performed a binary classification (apnea or normal) based on nasal airflow analysis using a CNN1D classifier and a balanced dataset. The network consists of three convolutional layers, each with 30 filters, and the size of kernel is [5, 1], the step size is 5, each filter is followed by a maximum pooling layer with a size of [2, 1], and a fully connected layer with a softmax activation function. By evaluating other activation functions, the author chose the activation function ReLU because it has the best accuracy and the fastest training time. Haider et al. [29] also tested CNN1D with three input signals using a hold-out method to analyze nasal airflow, abdominal respiration, and thoracic respiration signals, with 75% of the training and 25% of the test data set. Two back to back convolution layers and a subsampling layer (conv-conv-maxpooling) are used to establish a three-cascading state. However, the physiological signals used in their methods are inconvenient to measure, such as nasal pressure and airflow, which limit application scenarios. Our method can exceed their performance using only a single thoracic respiratory signal.

## 2.2. Logic rules in deep learning

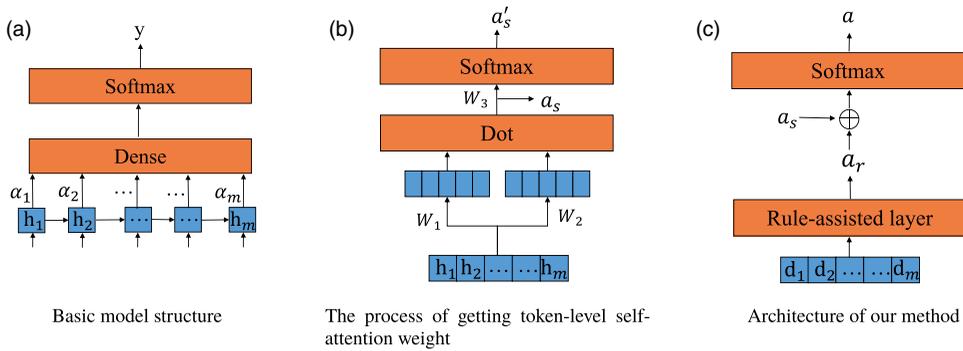
Logic rules embody high-level cognition and structured knowledge in the process of human communication. Incorporating rules into neural networks can be of great help to the learning process. The integration of common sense knowledge has also received a lot of attention in many tasks. Hu et al. [30] proposed a general framework that can use declarative first-order logic rules to improve a variety of neural networks. In particular, this paper developed a repeated knowledge distillation method that can transfer the structured information of logical rules to the weight of the neural network. The framework is implemented on the CNN network for sentence analysis and the RNN network for named entity recognition. Tandon et al. [31] proposed to use common sense knowledge as hard or soft constraints to bias the prediction of neural models for procedural text comprehension tasks. Xu et al. [32] used additional logic loss to enhance the training target as a means of applying soft constraints. The semantic loss used quantifies the probability of generating a satisfactory distribution by randomly sampling from the predicted distribution. Li et al. [3] proposed a framework that uses first-order logic to express knowledge without changing the end-to-end training method and integrates this structured knowledge into the neural network architecture. Our method extracts the key rules of the doctor's interpretation, introduce rule constraints into the neural network, and then use the rules that control attention to augment the network.

## 3. Our approach

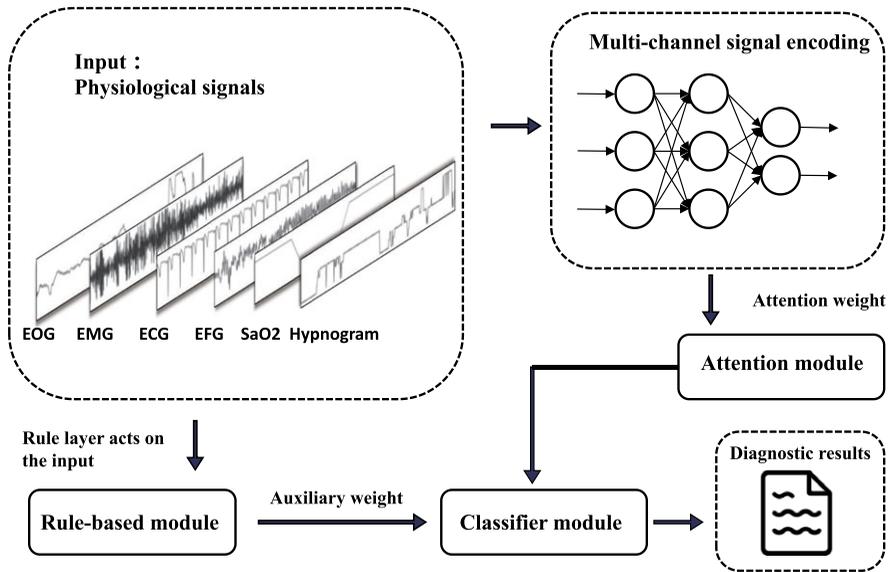
The architecture of our proposed method is shown in Fig. 1. The following introduction is divided three parts, including problem definition, multichannel model, and integration of rules. The whole process is shown in Fig. 2.

### 3.1. Problem definition

PSG contains a variety of physiological signals of patients, but the current research is limited to only a few of them. In addition to the commonly used signals, thoracic respiratory, abdominal respiratory, and nasal airflow, other signals are also related to the patient's sleep. Due to the different sampling rates of these signals, they have different dimensions. So we divide the PSG signals by sampling rate  $f_s$  to form multichannel data  $\mathcal{D} = \{\mathcal{D}^1, \mathcal{D}^2, \dots, \mathcal{D}^s\}$ ,  $s$  means the number of signal types by frequency. Now,  $\mathcal{D}^1 = \{\mathcal{D}^{11}, \mathcal{D}^{12}, \mathcal{D}^{13}, \mathcal{D}^{14}\}$  including EEG, ECG, EOG, and EMG,  $\mathcal{D}^2 = \{\mathcal{D}^{21}, \mathcal{D}^{22}, \mathcal{D}^{23}\}$  including thoracic respiratory, abdominal respiratory and nasal airflow and  $\mathcal{D}^3 = \{\mathcal{D}^{31}, \mathcal{D}^{32}\}$  including SpO2 and heart rate.  $\mathcal{D}^j = \{((\mathbf{x}_1, y_1), \mathbf{x}_2, y_2), \dots, (\mathbf{x}_j, y_j), \dots, (\mathbf{x}_n, y_n))\}$ ,  $\mathbf{x}_j \in \mathbb{R}^l$ ,  $l = f \cdot t$ ,  $t$  means the sampling time,  $y_i \in \{0, 1\}$



**Figure 1.** In our architecture, the initial input  $D$  is applied to the rule-assisted layer to obtain the auxiliary weight  $\alpha_r$ , and then  $\alpha_r$  is combined with the self-attention weight  $\alpha_s$  to obtain the final weight.



**Figure 2.** Integrating medical rules into models for apnea detection using PSG signals.

where 0 is normal and 1 is abnormal. Then, we use encode model  $E(\cdot; \theta)$  with different parameters to embedding these different dimensional segments into the same dimensional representations  $\mathbf{z}^{ch}$  for  $ch = 1, 2, \dots, k, k = |\mathcal{D}^1| + |\mathcal{D}^2| + \dots + |\mathcal{D}^k|$ . Now, given a special PSG signal segmentation  $\mathbf{d}^i \in \mathbb{R}^l$ , we can obtain a feature  $\mathbf{z}^i \in \mathbb{R}^m$  computed as  $E(\mathbf{d}^i; \theta^i)$  where  $m$  means the dimension of input after embedding. Then, we can use the same dimensional data  $Z = \{\mathbf{z}^1, \mathbf{z}^2, \dots, \mathbf{z}^k\}$  to train a classification model  $M(\cdot; \tilde{\theta})$  for diagnosis sleep apnea disease.

Clearly, the predictive capability of such model is limited because the amount of medical real labeling data is limited, especially in abnormal samples, which leads to weak model generalization ability. We propose a method of integrating medical interpretation rules into LSTM neural network with multichannel respiratory signals as input based on self-attention mechanism. First, we process the above features  $Z \in \mathbb{R}^{m \times k}$  by attention layer  $\text{Att}()$  to get the attention weights  $\alpha_s$ . Then, we build an auxiliary layer  $\text{Rule}()$  by medical rules to get auxiliary weights  $\alpha_r$ . Finally, the two parameters are connected by a real number parameter.

The following sections will describe how the above models can be computed in detail.

### 3.2. Multi-channel model

For data  $\mathcal{D} = \{\mathcal{D}^1, \mathcal{D}^2, \dots, \mathcal{D}^s\}$  of different frequencies, we encode the data  $\mathcal{D}$  separately to the same dimension using LSTM with different parameters. Formally,

$$\mathbf{z}^i = E(\mathbf{d}^i; \theta^i), i = 1, 2, \dots, k \tag{1}$$

where  $\mathbf{z}^i \in \mathbb{R}^{m \times 1}$  denotes the features of the same dimension after encoding,  $m$  means the dimension after encoding,  $E(\cdot; \cdot)$  represents the embedded model for the  $i$ -th signal,  $\mathbf{d}^i$  denotes the  $i$ th signal,  $k$  denotes the number of signal, and  $\theta^i$  denotes the parameters corresponding to each model. Then we get the next input  $X = \{\mathbf{z}^1, \mathbf{z}^2, \dots, \mathbf{z}^{ch}\}, X \in \mathbb{R}^{m \times k}$  to the subsequent classifier.

Here we have a feature  $X \in \mathbb{R}^{m \times k}$  as input to classifier.  $m$  means the dimension after encoding and  $k$  means the number of channel. We choose LSTM as the base model because LSTM neural networks is suitable for modeling sequence data. LSTM is an improved recurrent neural network (RNN) that can solve the problem that RNN cannot handle long-distance dependence. The hidden layer of the original RNN has only one state  $h$ , which is very sensitive to short-term inputs. The LSTM adds one state  $c$  and lets it save the long-term state, called cell state:

$$\mathbf{h}_t = LSTM(\mathbf{x}_t, \mathbf{h}_{t-1}). \tag{2}$$

Here,  $\mathbf{h}_t$  represents hidden state at time  $t$ . At time  $t$ , there are three inputs to the LSTM: the input value  $\mathbf{x}_t$  of the network at the current moment, the output value  $\mathbf{h}_{t-1}$  of the LSTM at the previous moment, and the state  $\mathbf{c}_{t-1}$  of the cell at the previous moment. There are two outputs of LSTM: the output value  $\mathbf{h}_t$  of the LSTM at the current moment and the state  $\mathbf{c}_t$  of the cell at the current moment. Formally,

$$\begin{aligned} \mathbf{f}_t &= \sigma(W_f \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + b_f) \\ \mathbf{i}_t &= \sigma(W_i \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + b_i) \\ \tilde{\mathbf{c}}_t &= \tanh(W_c \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + b_c) \\ \mathbf{c}_t &= \mathbf{f}_t \cdot \mathbf{c}_{t-1} + \mathbf{i}_t \cdot \tilde{\mathbf{c}}_t \\ \mathbf{o}_t &= \sigma(W_o \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + b_o) \\ \mathbf{h}_t &= \mathbf{o}_t \cdot \tanh(\mathbf{c}_t) \end{aligned} \tag{3}$$

where  $\sigma$  is a logical sigmoid function,  $\tanh$  is an activation function,  $W$  represents the weight matrix,  $b$  represents the bias term, and  $[\mathbf{h}_{t-1}, \mathbf{x}_t]$  represents a concatenation operation with  $\mathbf{h}_{t-1}$  and  $\mathbf{x}_t$ . The forget gate  $\mathbf{f}_t$  determines how much of the cell state  $\mathbf{c}_{t-1}$  from the previous moment is retained to the current state  $\mathbf{c}_t$ . The input gate  $\mathbf{i}_t$  determines how much of the input  $\mathbf{x}_t$  of the neural network at the current moment is saved to the cell state  $\mathbf{c}_t$ .  $\tilde{\mathbf{c}}_t$  is a new candidate vector created by the tanh layer and is added to the next cell state. The output gate  $\mathbf{o}_t$  controls how much of the cell state  $\mathbf{c}_t$  is output to the current output value  $\mathbf{h}_t$  of the LSTM. Now we integrate all the hidden state vectors into a matrix  $H$ .  $H \in \mathbb{R}^{u \times k}$ ,  $u$  means the length of hidden status.

### 3.3. Integration of rules

This section describes the integration of medical rules into the model based on the multichannel model described above, and this section includes token-level self-attention in LSTM, rule-assisted layer, and combination of weights.

#### 3.3.1. Token-level self-attention

Next, we take  $H$  as the input and use the dot-product attention mechanism to get attention weight. For easier integration with subsequent output of rule-assisted layer, we need token-level attention  $\alpha_s$ . To get the token-level attention weights, the weights are multiplied by a parameter vector after getting the dot product attention weights. The computational process is as follows:

$$V_{ij} = \sum_{q=1}^k (W_1 H)_{iq} (W_2 H)_{qj}^T, i, j = 1, 2, \dots, k, \tag{4}$$

$$(\alpha_s)_j = \sum_{q=1}^k (\mathbf{w}_3)_q V_{qj}, j = 1, 2, \dots, k. \tag{5}$$

Here,  $W_1, W_2$  is a weight matrix with a shape of  $k$  by  $u$ ,  $\mathbf{w}_3$  is a vector of parameter with size  $k$ ,  $V$  is the intermediate result, a matrix of similar weights, and  $\alpha_s$  is attention weight for each token with a size of  $m$ .

### 3.3.2. Rule-assisted layer

The American Academy of Sleep Medicine (AASM) has developed manual [33] for scoring of sleep and related event. The manual provides instructions for scoring sleep stages, respiratory events, and other sleep-related parameters to improve the accuracy and reproducibility of PSG measurements. The key medical rules for detecting sleep apnea events can be described as

(1) *There is a drop in the peak signal excursion by  $\geq 90\%$  of pre-event baseline using an oronasal thermal sensor (diagnostic study), positive airway pressure device flow (titration study), or an alternative apnea sensor.* (2) *The duration of the  $\geq 90\%$  drop in sensor signal is  $\geq 10$  s.*

We will borrow the predicate symbols defined in the natural language processing task. We define two rules to assist and constrain attention: (1)  $K_i \rightarrow A_i$  (2)  $R_i \wedge A_i \rightarrow A'_i$ .  $K_i$  denotes the relatedness,  $R_i$  denotes the weight after applying the rule to the original input,  $A_i$  denotes the attention weight obtained based on the internal relatedness, and  $A'_i$  denotes the weight after auxiliary and restriction.

The abnormal respiratory events that will be considered in the diagnosis of SA include apnea and hypopnea. The above rules are for detecting apnea. The difference between hypopnea and apnea lies in the degree of decline. The recommended hypopnea definition requires a 30% or greater drop in flow for 10 s or longer associated with  $\geq 4\%$  oxygen desaturation. This value of the drop is set as a hyperparameter  $\beta$ , and then BO is used to find the best value.

We extract key medical rules as additional knowledge to assist attention weights. Formally,

$$\alpha_r = \frac{1}{m} \sum_{i=1}^m \text{Rule}(\mathbf{d}^i) \tag{6}$$

where  $\mathbf{d}^i \in \mathcal{D}$ , the detailed process of Rule() is shown in Algorithm 1. We first label each segmentation with the corresponding baseline value using the annotation of the dataset based on each segmentation to obtain the baseline value closest to the corresponding time period.  $p_n$  represents the normal amplitude of breathing, which is the baseline value.  $p_c$  represents the signal amplitude of the current period.  $cnt$  represents number of slices that are continuously less than the baseline value.

### 3.3.3. Weight combination

Our purpose is to assist in modifying the attention weight through the restriction of the rule-assisted layer and combine the two in the following way:

$$\alpha = \text{softmax}(\alpha_s + \lambda \alpha_r), \tag{7}$$

$$H_r = \alpha \cdot H. \tag{8}$$

Here,  $\lambda$  is a non-negative hyperparameter. This hyperparameter determines the degree of restriction of the rule-assisted layer. *softmax*() ensures that the sum of all calculated weights is 1. The new matrix  $H_r$  is obtained by multiplying the weight vector  $\alpha$  and hidden state  $\mathbf{h}_i$ .  $H_r$  replaces  $H$  as the input of the subsequent fully connected layer. The loss function is the binary crossentropy as defined by

**Algorithm 1.** Rule-assisted layer

---

```

Input: input  $\mathcal{D}$ 
Output: auxiliary weight  $\alpha_r$ 
1 initialization:  $\alpha_r = 0, p_c, p_n = 0, cnt = 0, hyperparameter \tilde{\alpha} = 1.0;$ 
2 for  $j \leftarrow 1$  to  $len(\mathcal{D})$  do
3    $s = \mathcal{D}[j, :, :];$ 
4   for  $i \leftarrow 1$  to  $s.size[1]$  do
5     if  $p_c < \tilde{\alpha} \times p_n$  then
6        $cnt + = 1;$ 
7        $\alpha_r[j : i] = 1;$ 
8       if  $cnt > 2$  then break ;
9       else continue;
10    end
11    else
12       $cnt = 0$ 
13    end
14  end
15 end

```

---

$$L = \sum_{i=1}^N (y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)) \quad (9)$$

where  $N$  represents the number of samples for an epoch,  $y_i$  represents the true binary label of sample  $i$ , and  $\hat{y}_i$  represents the predicted probability of sample  $i$ .

## 4. Experiments and results

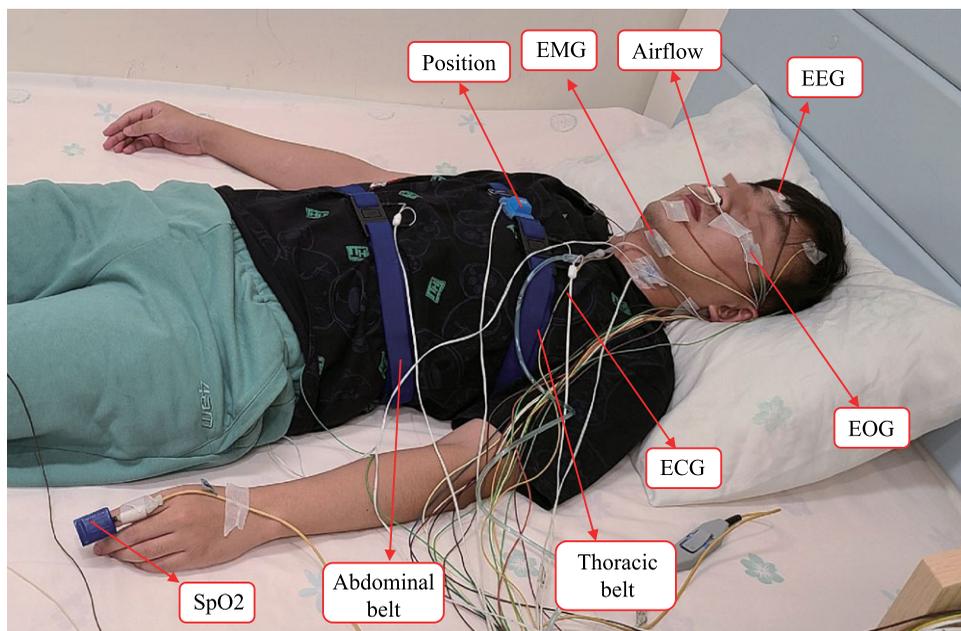
### 4.1. Data description

The Sleep Heart Health Study (SHHS)<sup>1</sup> [34, 35] is a multicenter cohort study implemented by the National Heart Lung & Blood Institute to determine the cardiovascular and other consequences of sleep-disordered breathing. The SHHS Visit 1 (SHHS-1) dataset represents data from the baseline and first follow-up visits, collected on 6441 individuals between 1995 and 1998. A sample of participants who met the inclusion criteria (age 40 years or older; no history of treatment of sleep apnea; no tracheostomy; no current home oxygen therapy) was invited to participate in the baseline examination of the SHHS, which included an initial polysomnogram. Polysomnograms were obtained in an unattended setting by trained and certified technicians. The recording consisted of: electroencephalogram (EEG), electrocardiogram (ECG), electrooculograms (EOG), electromyogram (EMG), thoracic respiration (TR) and abdominal respiration (AR), nasal airflow (NA), pulse oxygen saturation (SpO<sub>2</sub>), heart rate (HR), body position and ambient light as shown in the Fig. 3. Each recording has a signal file, event scoring, and epoch staging annotations.

### 4.2. Data processing

The raw physiological signal contains a wide range of noise due to subject motion, electrical interference, measurement noise, and other disturbances. Noise reduction methods are essential and frequently used in any sleep apnea detection method. To extract relevant respiratory information and reduce noise, the physiological respiratory signal is passed through a fourth-order low-pass Butterworth filter with a cutoff frequency of 0.7 Hz [36]. This cutoff frequency is chosen to preserve the main respiratory components while eliminating as much noise as possible [37]. Taking into account, the length of the apnea time in the data set and the doctor's recommendation, the signal is divided into 100 s epochs with a step of 1 s

<sup>1</sup><https://www.sleepdata.org/datasets/shhs>



**Figure 3.** A demonstration of SA diagnosis using polysomnography (PSG).

between them and adopts its original frequency. The sample is labeled according to the annotation file provided in the SHHS dataset. Then, we reduced the number of normal samples to approximately the same as the abnormal samples.

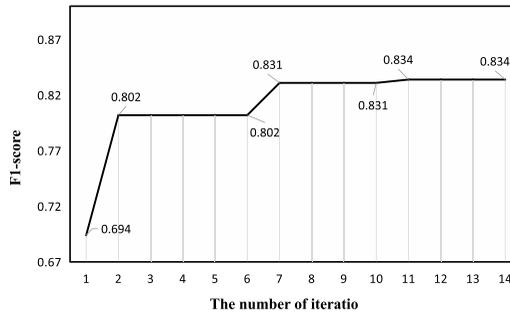
### 4.3. Experiment setup

We use LSTM as the basic model for classification, define the step size in LSTM as 4 s, and train an LSTM with a length of 25 given an observation window of 100 s. The LSTM network architecture is as follows: it consists of an LSTM layer and a dropout layer. The function of the dropout layer is to improve the generalization ability of the network to unknown data. Then, a dense layer with the relu activation function is added followed by a dropout layer. Finally, a dense layer with softmax activation function is added. The output produced by this activation function can be interpreted as the probability that the input epoch contains apnea. During training, the time step of the sample is set to 40 depending on the body's breathing cycle, so the shape of input reshapes to  $b \times t \times m$ ,  $b$  means batch size and  $t$  means time step. The ratio of the three in the train set, validation set, and test set is set to 5 : 2 : 3. The test set of all methods remains the same. We use optimization algorithm for stochastic gradient descent as the optimizer, specify a batch size of 128, an epoch of 100, and a learning rate of 0.001.

In this work, the proposed all models are implemented on Tensorflow and Keras libraries and simulated using a PowerLeader PR4908P server configured with  $8 \times 32\text{GB}$  RAM, Intel(R) Xeon(R) Gold 6154 CPU, and TITAN XP GPU.

We will evaluate the performance of the proposed methods and compare it with its counterpart. We use vanilla LSTM as the basic model, denoted as vLSTM. The vLSTM model with token-level self-attention mechanism is denoted as sLSTM. The sLSTM model with the rule-assisted layer is denoted as rLSTM.

The performance of the models is evaluated according to the following test criteria: accuracy  $Acc = (TP + TN)/(TP + TN + FP + FN)$ , precision  $Pre = TP/(TP + FP)$ , recall  $Rec = TP/(TP + FN)$ , and f1-score  $F1 = 2(Pre \times Rec)/(Pre + Rec)$ , where TP, TN, FP, and FN represent true-positive, true-negative, false-positive, and false-negative predictions, respectively.



**Figure 4.** Performance of parameters selection using Bayesian optimization.

#### 4.4. Result analysis

##### 4.4.1. Performance of multichannel model

In this section, we will compare the experimental effects of different signals as inputs. The inputs are divided into single-signal and multichannel signals. Single signal includes EEG, ECG, EOG, EMG, SpO<sub>2</sub>, HR, TR, AR, and NR. The multichannel signal includes three physician-recommended signals (TR, AR and NR) and PSG signals (all the above single signals). Given the same respiratory signal segmentations and the corresponding test set labels, we measure their prediction performance (i.e., accuracy, precision, recall, and f1-score).

In order to optimize the introduced two hyperparameters  $\lambda$  and  $\beta$ ,  $\lambda$  is a non-negative hyperparameter. This hyperparameter determines the degree of restriction of the rule-assisted layer.  $\beta$  represents the amplitude of the signal drop. We first use Bayesian optimization to automatically select the desired hyperparameters. Then, we will use the optimal parameters to build subsequent models. The result of Bayesian optimization is shown in Fig. 4. The higher the performance evaluation of the best candidate, the better the hyperparameter performance of the group. After Bayesian optimization, the hyperparameters we choose are  $\lambda = 0.5$ ,  $\beta = 0.8$ .

We did experiments with multiple signals as multichannel inputs to verify the effect of multidimensional data on the detection effect. As shown in Table I, it can be seen that the physician's suggested signal is superior to the other signals from the experimental results, and the performance of nasal airflow is the best in the single signal experiment. The results of multichannel models are overall better than those of single signal models, and the PSG signals with more signals are better. It can be seen that the multichannel model has some improvement in the overall.

##### 4.4.2. Performance of rule-assisted layer

In this section, we compare the proposed methods with two popular sleep apnea detection algorithms and a rule-based method. Given the same respiratory signal segmentations and the corresponding test set labels, we measure their prediction performance (i.e., accuracy, precision, recall, and f1-score).

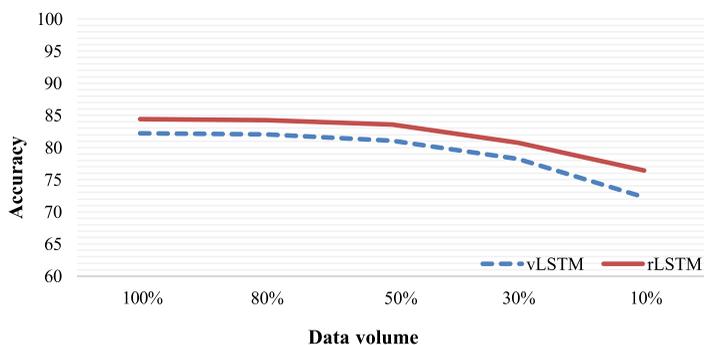
Next, our proposed method is compared with the existing methods together. The comparison algorithm uses the same data for training. As shown in Table II, the performance of our basic model vLSTM is slightly better than CNN1D, which is the best existing method in terms of accuracy, f1-score, and precision. The sLSTM model that introduces the token-level self-attention mechanism has a certain improvement compared to vLSTM, which shows that the self-attention mechanism can help improve performance. After adding the rule-assisted layer to assist the attention weight, the model rLSTM has a slight decrease in precision compared to sLSTM, but it has a certain improvement in the other three evaluation metrics, especially in accuracy. With the additional domain knowledge, the performance of the proposed rLSTM method is comparable in all evaluation metrics compared to the best prediction performance of existing methods. The average degradation on recall is only 0.0322, but the average improvement is 0.0326, 0.0703, and 0.0178 on accuracy, precision, and f1-score, respectively.

**Table I.** Comparison of single signal and multiple signals: EEG, ECG, EOG, EMG, SpO<sub>2</sub>, heart rate (HR), thoracic respiration (TR), abdominal respiration (AR), nasal airflow (NA), three physician-recommended signals (TPRS).

Signal	Accuracy	Recall	Precision	F1-score
EEG	74.32	74.29	75.22	74.75
ECG	75.19	74.85	76.11	75.47
EOG	74.29	73.54	75.17	74.35
EMG	73.89	73.22	74.4	73.81
SpO <sub>2</sub>	72.29	72.01	72.87	72.44
HR	69.52	68.41	70.65	69.51
TR	78.41	78.24	79.05	78.64
AR	80.24	79.78	81.01	80.39
NA	80.82	<b>80.49</b>	81.48	80.98
TPRS	81.01	80.14	81.22	80.68
PSG	<b>81.35</b>	77.91	<b>85.47</b>	<b>81.52</b>

**Table II.** Comparison with existing models.

Method	Accuracy	Recall	Precision	F1-score
LSTM [9]	75.01	67.85	76.58	71.95
CNN1D [18]	81.16	<b>83.51</b>	79.18	81.28
Rule-based	76.72	83.44	72.31	77.56
vLSTM	81.35	77.91	85.47	81.52
sLSTM	82.21	78.14	86.16	81.95
rLSTM	<b>84.89</b>	80.57	<b>86.21</b>	<b>83.29</b>



**Figure 5.** The impact of data volume.

#### 4.4.3. Impact of data volume

To verify whether the rule layer helps to alleviate the need for data, we choose models sLSTM and rLSTM for comparison experiments. We train the models using 100%, 80%, 50%, 30%, and 10% of the training data, respectively, and then validate the models using the same test set. As shown in Fig. 5, the overall performance of rLSTM is better than that of sLSTM. As the amount of data decreases, the overall trend of the two models is decreasing, but it can be seen that the decline of the rLSTM model has a certain degree of relaxation compared with the sLSTM model. This shows that additional domain knowledge can play a certain role in alleviating the need for data.

## 5. Conclusion

In this paper, we propose a new method to extract key rules in sleep apnea detection as additional domain knowledge to assist and constrain attention weights to improve the generalization ability of the model and alleviate the need for data. Compared with the current state-of-the-art method, the results of evaluating the model in the same public data set show a considerable improvement. With the additional domain knowledge, the performance of the proposed method is comparable in all evaluation metrics compared to the best prediction performance of existing methods. The average degradation on recall is only 0.0322, but the average improvement is 0.0326, 0.0703, and 0.0178 on accuracy, precision, and f1-score, respectively. Our models can benefit from additional external domain knowledge during training and inference, especially in the case of limited training data.

**Author contributions.** Jianqiang Li proposed the methodology in this work. Xiaoxiao Song completed the experiments and the draft. Yanning Lin helped with supplementary experiments and paper revising. Junya Wang processed the datasets for experiments. Dongying Guo provided professional support in healthcare domain. Jie Chen guided the progress of this research work. All authors have worked proportionally and given approval to the present research.

**Financial support.** This work was supported in part by the National Key R&D Program of China under Grant 2020YFA0908700, in part by the National Nature Science Foundation of China under Grants 62072315, 62073225, and 61836005, in part by the Shenzhen Science and Technology Program under Grant JCYJ20210324093808021 and Grant JCYJ20220531102817040, in part by the Natural Science Foundation of Guangdong Province-Outstanding Youth Program under Grant 2019B151502018, in part by the Guangdong “Pearl River Talent Recruitment Program” under Grant 2019ZT08X603, in part by the Shenzhen Science and Technology Innovation Commission under Grant R2020A045.

**Conflicts of interest.** The authors declare no conflicts of interest.

**Ethical standards.** Not applicable.

## References

- [1] M. J. Sateia, “International classification of sleep disorders,” *Chest* **146**(5), 1387–1394 (2014).
- [2] Z. Li, C. Deng and K. Zhao, “Human-cooperative control of a wearable walking exoskeleton for enhancing climbing stair activities,” *IEEE Trans. Ind. Electron.* **67**(4), 3086–3095 (2019).
- [3] T. Li and V. Srikumar, “Augmenting neural networks with first-order logic (2019), arXiv preprint [arXiv:1906.06298](https://arxiv.org/abs/1906.06298).
- [4] H. Watson, M. A. Sackner and A. S. Belsito, Method and apparatus for distinguishing central obstructive and mixed apneas by external monitoring devices which measure rib cage and abdominal compartmental excursions during respiration, Oct. 18 1988, US Patent 4,777,962.
- [5] X. Wu and Z. Li, “Cooperative manipulation of wearable dual-arm exoskeletons using force communication between partners,” *IEEE Trans. Ind. Electron.* **67**(8), 6629–6638 (2019).
- [6] J. Zhang, Q. Zhang, Y. Wang and C. Qiu, A Real-Time Auto-Adjustable Smart Pillow System for Sleep Apnea Detection and Treatment. In: *Proceedings of the IPSN* (IEEE, 2013) pp. 179–190.
- [7] T. Young, M. Palta, J. Dempsey, J. Skatrud, S. Weber and S. Badr, “The occurrence of sleep-disordered breathing among middle-aged adults,” *New Engl J. Med.* **328**(17), 1230–1235 (1993).
- [8] T. Young, L. Evans, L. Finn and M. Palta, “Estimation of the clinically diagnosed proportion of sleep apnea syndrome in middle-aged men and women,” *Sleep* **20**(9), 705–706 (1997).
- [9] Z. Li, C. Xu, Q. Wei, C. Shi and C.-Y. Su, “Human-inspired control of dual-arm exoskeleton robots with force and impedance adaptation,” *IEEE Trans. Syst. Man Cybern. Syst.* **50**(12), 5296–5305 (2018).
- [10] T. Gislason and B. Benediktsdottir, “Snoring, apneic episodes, and nocturnal hypoxemia among children 6 months to 6 years old: An epidemiologic study of lower limit of prevalence,” *Chest* **107**(4), 963–966 (1995).
- [11] S. Ancofi-Israel, E. R. DuHamel, C. Stepnowsky, R. Engler, M. Cohen-Zion and M. Marler, “The relationship between congestive heart failure, sleep apnea, and mortality in older men,” *Chest* **124**(4), 1400–1405 (2003).
- [12] P. de Chazal, T. Penzel and C. Heneghan, “Automated detection of obstructive sleep apnoea at different time scales using the electrocardiogram,” *Physiol. Meas.* **25**(4), 967–983 (2004).
- [13] R. Agarwal and J. Gotman, “Computer-assisted sleep staging,” *IEEE Trans. Biomed. Eng.* **48**(12), 1412–1423 (2001).
- [14] G. Li, Z. Li and Z. Kan, “Assimilation control of a robotic exoskeleton for physical human-robot interaction,” *IEEE Robot. Autom. Lett.* **7**(2), 2977–2984 (2022).
- [15] X. Ren, Y. Liu, Y. Hu and Z. Li, “Integrated task sensing and whole body control for mobile manipulation with series elastic actuators,” *IEEE Trans. Autom. Sci. Eng.* **20**(1), 413–424 (2022).

- [16] T. Zhao, Y. Liu, Z. Li, C.-Y. Su and Y. Feng, “Adaptive control and optimization of mobile manipulation subject to input saturation and switching constraints,” *IEEE Trans. Autom. Sci. Eng.* **16**(4), 1543–1555 (2018).
- [17] H. Su, Y. Hu, H. R. Karimi, A. Knoll, G. Ferrigno and E. De Momi, “Improved recurrent neural network-based manipulator control with remote center of motion constraints: Experimental results,” *Neural Networks* **131**, 291–299 (2020).
- [18] H. Su, W. Qi, Y. Schmirander, S. E. Ovrur, S. Cai and X. Xiong, “A human activity-aware shared control solution for medical human–robot interaction,” *Assembly Autom.* **42**(3) (2022), no. ahead-of-print.
- [19] Y. Xu, H. Su, G. Ma and X. Liu, “A novel dual-modal emotion recognition algorithm with fusing hybrid features of audio signal and speech context,” *Complex Intell. Syst.* (2022).
- [20] D. Wang, G. Ma and X. Liu, “An intelligent recognition framework of access control system with anti-spoofing function,” *Aims Math.* **7**(6), 10495–10512 (2022).
- [21] Z. Li, K. Zhao, L. Zhang, X. Wu, T. Zhang, Q. Li, X. Li and C.-Y. Su, “Human-in-the-loop control of a wearable lower limb exoskeleton for stable dynamic walking,” *IEEE/ASME Trans. Mech.* **26**(5), 2700–2711 (2020).
- [22] X. Liu, W. Jiang, H. Su, W. Qi and S. S. Ge, “A control strategy of robot eye-head coordinated gaze behavior achieved for minimized neural transmission noise,” *IEEE/ASME Trans. Mech.* **28**(2), 956–966 (2022).
- [23] M. Zhou, X. Wang, J. Weiss, A. Eslami, K. Huang, M. Maier, C. P. Lohmann, N. Navab, A. Knoll and M. A. Nasseri. Needle Localization for Robot-Assisted Subretinal Injection Based on Deep Learning. **In:** *Proceedings of ICRA* (IEEE, 2019) pp. 8727–8732.
- [24] X. Liu, X. Li, H. Su, Y. Zhao and S. S. Ge, “The opening workspace control strategy of a novel manipulator-driven emission source microscopy system,” *ISA Trans.* **134**, 573–587 (2022).
- [25] H. D. Park, Y. Han and J. H. Choi. Frequency-Aware Attention Based LSTM Networks for Cardiovascular Disease. **In:** *Proceedings of ICTC* (IEEE, 2018) pp. 1503–1505.
- [26] T. Van Steenkiste, W. Groenendaal, D. Deschrijver and T. Dhaene, “Automated sleep apnea detection in raw respiratory signals using long short-term memory neural networks,” *IEEE J. Biomed. Health* **23**(6), 2354–2364 (2018).
- [27] V. Thorey, A. B. Hernandez, P. J. Arnal and E. H. During. Ai vs Humans for the Diagnosis of Sleep Apnea. **In:** *Proceedings of EMBC* (IEEE, 2019) pp. 1596–1600.
- [28] R. Haidar, I. Koprinska and B. Jeffries. Sleep Apnea Event Detection from Nasal Airflow Using Convolutional Neural Networks. **In:** *Proceedings of ICONIP* (2017) pp. 819–827.
- [29] R. Haidar, S. McCloskey, I. Koprinska and B. Jeffries. Convolutional Neural Networks on Multiple Respiratory Channels to Detect Hypopnea and Obstructive Apnea Events. **In:** *Proceedings of IJCNN* (IEEE, 2018) pp. 1–7.
- [30] Z. Hu, X. Ma, Z. Liu, E. Hovy and E. Xing, “Harnessing deep neural networks with logic rules (2016), arXiv preprint [arXiv:1603.06318](https://arxiv.org/abs/1603.06318).
- [31] N. Tandon, B. D. Mishra, J. Grus, W.-t. Yih, A. Bosselut and P. Clark, “Reasoning about actions and state changes by injecting commonsense knowledge (2018), arXiv preprint [arXiv:1808.10012](https://arxiv.org/abs/1808.10012).
- [32] J. Xu, Z. Zhang, T. Friedmann, Y. Liang and G. Broeck. A Semantic Loss Function for Deep Learning with Symbolic Knowledge. **In:** *Proceedings of the ICML* (PMLR, 2018) pp. 5502–5511.
- [33] R. B. Berry, R. Budhiraja, D. J. Gottlieb, D. Gozal, C. Iber, V. K. Kapur, C. L. Marcus, R. Mehra, S. Parthasarathy, S. F. Quan, S. Redline, K. P. Strohl, S. L. Davidson Ward and M. M. Tangredi, “Rules for scoring respiratory events in sleep: Update of the 2007 aasm manual for the scoring of sleep and associated events: Deliberations of the sleep apnea definitions task force of the american academy of sleep medicine,” *J. Clin. Sleep Med.* **8**(5), 597–619 (2012).
- [34] S. F. Quan, B. V. Howard, C. Iber, J. P. Kiley, F. J. Nieto, G. T. O’Connor, D. M. Rapoport, S. Redline, J. Robbins, J. M. Samet and P. W. Wahl, “The sleep heart health study: Design, rationale, and methods,” *Sleep* **20**(12), 1077–1085 (1997).
- [35] G.-Q. Zhang, L. Cui, R. Mueller, S. Tao, M. Kim, M. Rueschman, S. Mariani, D. Mobley and S. Redline, “The national sleep research resource: Towards a sleep data commons,” *J. Am. Med. Inform. Assoc.* **25**(10), 1351–1358 (2018).
- [36] T. Van Steenkiste, W. Groenendaal, J. Ruyssinck, P. Dreesen, S. Klerkx, C. Smeets, R. de Francisco, D. Deschrijver and T. Dhaene. Systematic Comparison of Respiratory Signals for the Automated Detection of Sleep Apnea. **In:** *Proceedings of the EMBC* (IEEE, 2018) pp. 449–452.
- [37] D. A. Hettrick and T. M. Zielinski, “Bioimpedance in Cardiovascular Medicine,” **In:** *Encyclopedia of Medical Devices and Instrumentation* (2006).