Internet of Things 20 (2022) 100613

Contents lists available at ScienceDirect

Internet of Things

journal homepage: www.elsevier.com/locate/iot

Research article

Single-lead ECG based multiscale neural network for obstructive sleep apnea detection

Zhiya Wang [a,1], Caijing Peng [b], Baozhu Li [c,1], Thomas Penzel [d], Ran Liu [e], Yuan Zhang [a,*], Xinge Yu [f,g]

[a] College of Electronic and Information Engineering, Southwest University, Chongqing 400715, China
[b] Department of Pediatric Respiration, Chongqing Ninth People's Hospital, Chongqing 400700, China
[c] Internet of Things Smart City Innovation Platform, Zhuhai Fudan Innovation Institute, Zhuhai 519031, China
[d] Interdisciplinary Centre of Sleep Medicine, CC 12, University Hospital Charité Berlin, Berlin, Germany
[e] Department of Biomedical Engineering, School of Medicine, Tsinghua University, Beijing 100084, China
[f] Department of Biomedical Engineering, City University of Hong Kong, Hong Kong
[g] Hong Kong Centre for Cerebro-Cardiovascular Health Engineering (COCHE), Hong Kong Science Park, 999077, Hong Kong

ARTICLE INFO

Keywords:
Wearable ECG
Obstructive sleep apnea
Multi-scale neural network
Deep learning

ABSTRACT

Obstructive sleep apnea (OSA) is a common sleep disorder characterized by frequent cessation of breathing during sleep, which cannot be easily diagnosed at the early stage due to the complexity and labor intensity of the polysomnography (PSG). Using a ECG device for OSA detection provides a convenient solution in the current Internet of Things scenario. However, previous intelligent analysis algorithms mainly rely on single scale network, therefore the discriminative ECG representations cannot be identified, which affects the accuracy of OSA detection. We report a multiscale neural network URNet for OSA detection by optimizing the deep learning networks and integrating Unet with ResNet. The URNet automatically extracts delicate features from the RR interval of single-lead ECG and processes convolution blocks with different scales by skip connections, so that the network can fuse features collected from both shallow and deep levels. For each OSA segment identification, URNet achieves an accuracy of 90.4%, a sensitivity of 83.3%, a specificity of 94.8% and an F1 of 89.6% on the Apnea-ECG dataset. The result indicates that our approach provides major improvements compared to the state-of-the-art methods. The URNet model proposed in this study for unobstructive OSA detection has good potential application in daily sleep health.

1. Introduction

Sleep, as the basic physiological process of life, can eliminate fatigue, generate new vitality and improve immunity. Obstructive sleep apnea (OSA) is a chronic disorder caused by repeated upper-airway collapse during sleep, resulting in recurrent nocturnal asphyxia [1]. Patients with untreated OSA expose to increased risk of hypertension, stroke, heart failure, diabetes, car accidents, and depression [2]. Several studies have shown that the prevalence of OSA could be estimated to around 34% in men aged 30–70 years and 17% in women aged 30–70 years [3]. Although OSA is a severe chronic disease, it can be effectively treated as long as accurately diagnosed by doctors. Therefore, it is significant to raise public health awareness and develop smart sleep health technology for early and accurate OSA detection.

* Corresponding author.
E-mail address: yuanzhang@swu.edu.cn (Y. Zhang).
[1] Zhiya Wang and Baozhu Li contributed equally to this work.

https://doi.org/10.1016/j.iot.2022.100613
Received 20 July 2022; Received in revised form 28 August 2022; Accepted 5 September 2022
Available online 8 September 2022
2542-6605/

Usually, the screening and diagnosing of OSA require patients to wear at least 22 electrodes over night through the polysomnography (PSG) in a sleep center to obtain signals, such as electroencephalogram (EEG), electrocardiography (ECG), and electromyogram (EMG). The process of PSG collection is quite discomfort and time-consuming. These difficulties limit the application of PSG in daily life and may result in delay of OSA treatment.

In order to find an alternative to PSG, researchers proposed some lightweight wearable electronic devices to detect OSA through physiological signals, such as blood oxygen saturation (SpO2) [4,5], snoring [6], respiratory signals [7], ECG, etc. Clinical experiments have observed that the significant correlation between OSA and arrhythmia may lead to abnormal RR intervals [8]. It is desirable to obtain sleep breathing events from wearable ECG device [9,10] which is less disruptive during sleep. For this reason, we employ wearable ECG for the detection of OSA and compared our results to other studies with the same database.

The review [11] summarized computational models and methods for intelligent sleep medicine. Residual network (ResNet) can be applied to solve the degradation problem by introducing shortcut connection. The Unet evolved from a fully convolutional network (FCN), is characterized by its ability of feature fusion. It can stitch together features at different scales, and the last upsampling fuses the output of both the first convolutional module and the previous upsampling.

Based on our previous researches [12,13] and aiming at undisturbed early OSA detection, single-lead ECG signals are adopted for developing a novel URNet model. Our main contributions are as follows:

(1) We propose a single-lead ECG based multiscale neural network URNet for early OSA detection, which offers a promising alternative to the current methods. Compared with other state-of-the-art researches on the same dataset, it shows superior performance.

(2) In the feature extraction stage, a one-dimension ResNet is designed as the encoding module of the network to capture the subtle variations of the ECG signals at different layers. The skip connection connects the encoder to the decoder, and retains the effective features in the lost information at different layers of the encoder, so that the network could integrate multi-scale features to have stronger ability in identifying OSA events from ECG signals.

(3) We also implement real time experiment on a low-power embedded platform NVIDIA Jetson Xavier NX.

2. Related works

Early studies endeavored to solve manual feature extraction and to select suitable classifiers. Hwang et al. [4] introduced a method to automatically detect apnea and estimate the apnea–hypopnea (AHI) index by a regression model using the morphological features in the fluctuation of blood oxygen saturation. Heenam et al. [5] estimated the AHI index through a regression model by extracting quantitative characteristics caused by apnea during the changes of SpO2. P. Temrat et al. [6] extracted the time domain features of snoring signals after apnea events, and classified different types of snoring by leave-one-out cross-validation to distinguish the severity of OSA. Sharma et al. [14] used the error of QRS complex approximation and the coefficients of Hermite decomposition to simultaneously extract features and energy from RR intervals in ECG. Subsequently, they segmented the apnea and normal segments by four traditional classifiers.

As time goes by, CNN has become an advanced technology in image processing and computer vision [15]. Deep learning networks are able to identify the structure of lower level data and provide feature input for higher level data learning. Steenkiste et al. [7] presented a deep learning algorithm based on the improved LSTM model to automatically extract features and detect events in respiratory signals. Wang et al. [16] applied a modified LeNet-5 network to identify sleep apnea. Shen et al. [17] developed a multiscale network to extract RR interval information and utilized it for OSA segment identification. Li et al. [18] adopted a sparse autoencoding approach on feature representation of label-free ECG signals and employed a SVM classifier to detect OSA. Similarly, Feng et al. [19] proposed an unsupervised feature learning deep network model for frequency domain superposition SAE (FSSAE), and successfully introduced the unsupervised learning to comprehensive feature extraction of sleep apnea task. Nasifoglu et al. [20] proposed model generates scalogram and spectrogram by transforming preprocessed 30 s ECG segments from time domain to the frequency domain, and predicted OSA through them.

Although previous studies have achieved satisfactory performance, there are still problems that need to be solved. Existing machine learning methods mainly use the frequency domain, time domain and some non-linear features to build models, relying on researchers' prior knowledge and experience in the relevant field. Deep learning approaches have achieved notable achievements, as automatic feature extraction can solve some drawbacks or traditional machine learning methods. However they usually have two limitations. Firstly, some of them convert the one-dimensional ECG signal into a two-dimensional spectrogram, which increase the complexity of the preprocessing stage. Second, while most methods are mainly performed on a single scale network, they fail to identify the discriminative ECG representations, thus affecting the accuracy of OSA detection.

To solve these problems, we propose a novel approach to detect OSA based on single-lead ECG and multiscale neural network URNet. It aims to learn the multiscale features from the RR intervals to improve the performance of the classifier.

3. Method

The overall framework of URNet is shown in Fig. 1. Firstly, raw ECG signal is preprocessed and cut into segments to obtain the RR intervals (interval between R peaks) and amplitudes as input. Next, the extracted information is fed into the URNet. Then features of RR intervals are automatically extracted by four encoders, and the skip connection retains the information lost at different layers in the encoder module, so that the URNet can learn more effective multiscale features. Finally, the features extracted by the network are fed into the classifier for classification. Details on each module will be presented in Section 3.2.

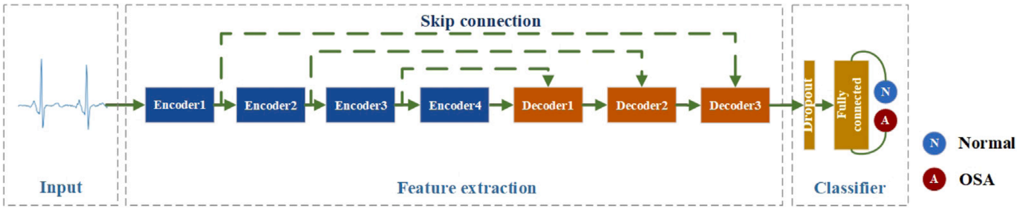


Fig. 1. The architecture of URNet.

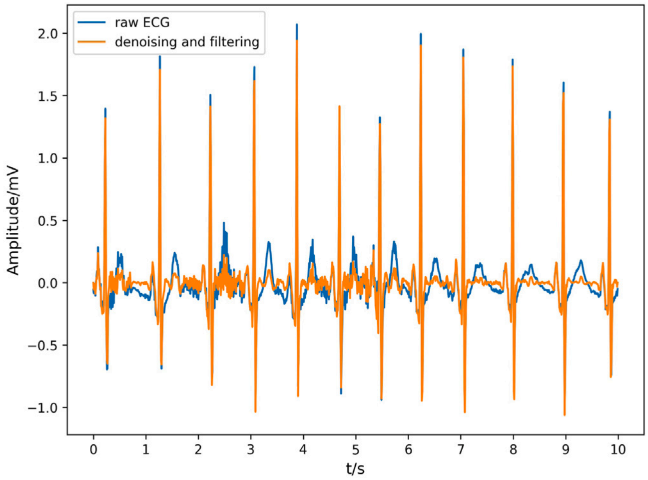


Fig. 2. Raw ECG vs. denoised ECG.

3.1. Preprocessing

During the process of ECG collecting throughout the night, baseline wanders occur due to human respiration or the minor displacement the electrode itself. In this study, Finite Impulse Response (FIR) band-pass (0.5–48 Hz) filter was applied to remove the noise and baseline wander. Subsequently, the RR intervals and amplitudes in the ECG signal were extracted as inputs to the model.

Several studies indicated that information of the adjacent segment assists the OSA detection of the current segment [21]. The accuracy is high when $t = 2$ during the experiment of the $\pm t$ surrounding segments before and after it(as shown in Fig. 8). Overlapping segmentations of five minutes are applied to the raw signal. Fig. 2 shows the 10-second raw ECG signal named "a01" in the Apnea-ECG training dataset, and the denoised signal after removing industrial frequency interference and baseline wanders.

After the preprocessing, the R peaks are identified by Hamilton algorithm, and the RR intervals and the amplitudes are extracted after correcting the R peaks. We choose the median filter proposed in [22] to remove some physiologically unexplainable points on the RR intervals. Since the input length of URNet must be the same, the revised RR intervals and the amplitudes are obtained by cubic interpolation.

3.2. URNet

URNet, evolved from ResNet and Unet, is a multiscale feature extraction based deep convolutional neural network. The architecture consisted of four encoders and three decoders for single-lead ECG based OSA detection. The filter number of the network, the connection structure of the encoder and decoder is shown in Fig. 3.

Fig. 4 shows the proposed encoder consisting of two operations: down-sampling module (DOWN_SAMP) and convolution module(CONV). They all utilize the shortcut principle proposed in ResNet. The two-dimension 3 × 3 convolution in ResNet-18 is modified into one-dimension with a kernel size of 5 to accommodate the detection of OSA from single-lead ECG signal. Both DOWN_SAMP and CONV modules, consists of two convolutional layers, two batch normalization (BN) layers and a ReLU activation function. ReLU can be expressed as:

$$g(x) = max(0, x) \tag{1}$$

where x represents the input data.

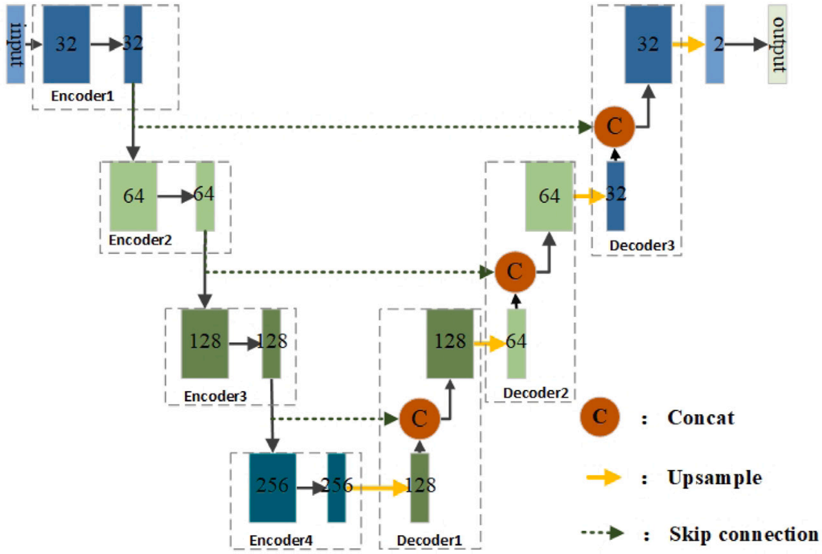


Fig. 3. The details of URNet.

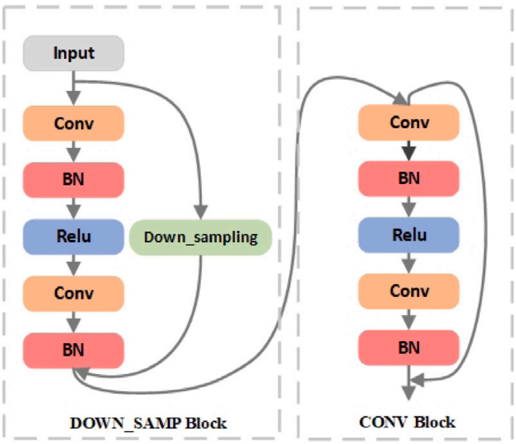


Fig. 4. Encoder module.

The convolutional layer is composed of several convolutional units and used to extract different features of the input. The latter BN layer, represented below enable the output to be close to the standard normal distribution, speed up the training speed, and avoid the disappearance of the gradient.

$$\hat{x}_i = \frac{x_i - E\left[x_i\right]}{\sqrt{\mathrm{Var}\left[x_i\right]}} \tag{2}$$

Here x_i is the batch input data, and $E[x_i]$ and $\sqrt{Var[x_i]}$ denote the average and standard deviation of training data, respectively.

Since the input feature map of DOWN_SAMP and CONV have different channels, an extra convolutional layer is added to the DOWN_SAMP. In summary, the encoder extracts multiscale feature in a hierarchical order, which is specially designed for classification tasks.

As shown in Fig. 5, the decoder contains three convolution layers and the first layer implements upsampling (up_samp) to recover the length of the signal. Decoder module takes the output of the encoder as input and applies transposed convolution to upsampling. Subsequently, two other convolutional layers and ReLu activation function are applied to make the decoder smoother. In order to make full use of all the features, the skip connections of the Unet are retained between the CONV module of each encoder and decoder (the dotted line indicates the skip connections). The network fuses multiscale characteristics by processing different dimensional features through convolution operation. In the end, the classifier employs the full connection layer to predict each signal segment.

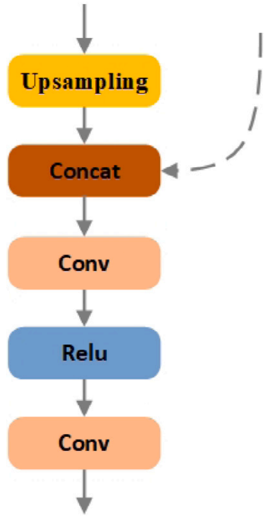


Fig. 5. Decoder module.

Table 1
Detailed network parameters.

Stage	Layer	Parameter
Encoder(DOWN_SAMP 1)	Conv1 Conv2 Downsample	K = 5, S = 2, pad = 2 K = 5, S = 1, pad = 2 K = 1, S = 5, pad = 0
Encoder(DOWN_SAMP 2)	Conv1 Conv2 Downsample	K = 5, S = 4, pad = 2 K = 5, S = 2, pad = 2 K = 1, S = 4, pad = 0
Encoder (CONV)	Conv1 Conv2	K = 5, S = 1, pad = 2 K = 5, S = 1, pad = 2
Decoder	up_samp Conv1 Conv1	K = 4/(5), S = 4/(5), pad = 0 K = 5, S = 1, pad = 2 K = 5, S = 1, pad = 2

Table 1 presents the detailed parameters of URNet. The K, S and pad represent the kernel size, strides and padding, respectively. The Encoder (DOWN_SAMP 1) in the table refers to the parameters of the DOWN_SAMP block in the first three encoders, while the Encoder (DOWN_SAMP 2) refers to the parameters of the DOWN_SAMP block in the fourth encoder. All the parameters of the CONV block are the same. The up_samp of the decoder is to restore the length, so that the parameters of the first decoder are K = 4, S = 4 and pad = 0, and the last two decoders are K = 5, S = 5 and pad = 0.

4. Experiments and results

4.1. Dataset

4.1.1. Apnea-ECG dataset

The dataset was provided by Phillips University [23]. It contains 70 single-lead ECG recordings, divided into training set (records named a01 through a20, b01 through b05 and c01 through c10) and withheld set (records named x01 through x35), with sampling frequency of 100 Hz and duration between 6.7 and 9.6 h. The technician annotated each 1-minute ECG recording in combination with other signals in the PSG (marked as A if an OSA event occurred within this minute; otherwise, marked as N). But this dataset did not distinguish between categories of respiratory events (obstructive and mixed apnea events were uniformly labeled A). This study uses the release set as the training set and the withheld set as a testing set. The total segments processed in our algorithm include 16 709 training samples (Normal : 10 236, OSA : 6473) and 16 945 testing samples (Normal : 10 455, OSA : 6490).

4.1.2. Local dataset

The study design was approved by the ethics committee of the Chongqing Ninth People's Hospital according to the ethics research review code 2021-SCI-004. The dataset recorded overnight PSG recordings of 49 male and 13 female patients aged 24–77 years, sampled at 256 Hz. Table 2 shows the demographics of the dataset. Unlike the Apnea-ECG Dataset, this dataset annotates the onset time and duration of respiratory events (obstructive, central, mixed apnea and hypopnea, and periodic respiratory episodes).

Table 2
The demographics of the local dataset.

Value	BMI (kg/m^2)	AHI (per h)	TRT (min)	TST (min)
Maximum	41.8	88.1	567.4	549.5
Minimum	21.7	5.3	358	261
Mean	28.6	31.8	495.7	452.6
SD	4.56	23.08	38.32	48.78

BMI: body mass index; AHI: apnea–hypopnea index; TRT: total recording time; TST: total sleep time; SD: standard deviation.

Considering that our study focus on OSA detection by the segment, the entire ECG recording was cut into 5-minute intervals. By apnea definition [24], an event lasts at least 10 s. Accordingly in this study, if apnea or hypopnea continued for 10 s or more, that minute is identified as an apnea. After the data preprocessing, the 62 records were randomly divided into the training set (52 subjects; Normal: 14 586, OSA: 8328) and the testing set (10 subjects; Normal: 2565, OSA: 2009).

4.2. *Experimental settings*

The experiment utilized grid search to find the optimal hyperparameters. We set the total number of training epochs to 100. The batch size was set as 128 to ensure that the model converges and does not over-fit. At the end of each epoch, we evaluated the testing set to save the results. The learning rate was 0.0003 by optimizing the Adam update rule [23]. The cross entropy function employed in the URNet is defined as:

$$L = -\frac{1}{n}\sum_{i=1}^{n}\left[p_i \cdot \log\left(y_i\right) + \left(1 - p_i\right) \cdot \log\left(1 - y_i\right)\right] \tag{3}$$

where n denotes the number of training samples, and p_i and y_i denotes the predicted and true labels of the samples, respectively.

4.2.1. *Offline analysis*

The data preprocessing and feature extraction are implemented in the PyCharm 2020.2.2 environment on an Intel(R) Core(TM) i7-8700K CPU @ 3.70 GHz with 64 GB RAM operating system. The hyperparameter optimization, training and test experiments of the URNet were run on a server with a GTX1080Ti based on the Torch framework.

4.2.2. *Embedded online analysis on NVIDIA Jetson Xavier NX*

NVIDIA Jetson Xavier NX is a device that can run deep learning algorithm with low power and very low consumption (10 W). It is suitable for practical application, such as design a smart thermography camera to diagnose the electrical equipment [25] and real time semantic segmentation [26]. The GPU is based on an NVIDIA Volta architecture with 384 CUDA cores and 48 Tensor cores. Our algorithm has been ported from an offline Python implementation to an online and embedded one to satisfy the E-health applications.

4.3. *Evaluation*

The model is estimated by accuracy, sensitivity, specificity, F1 value and confusion matrix.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

where TP, TN, FP, and FN stand for "true positives", "true negatives", "false positives" and "false negatives", respectively.

$$Sensitivity = \frac{TP}{TP + FN} \tag{5}$$

$$Specificity = \frac{TN}{TN + FP} \tag{6}$$

$$F1 = \frac{2 * TP}{2 * TP + FN + FP} \tag{7}$$

$$Confusionmatrix = \begin{bmatrix} TN & FP \\ FN & TP \end{bmatrix} \tag{8}$$

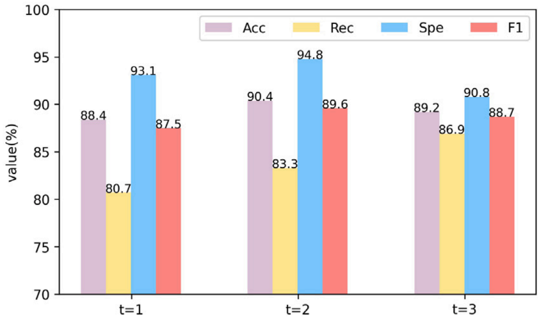


Fig. 6. The results of different periods t.

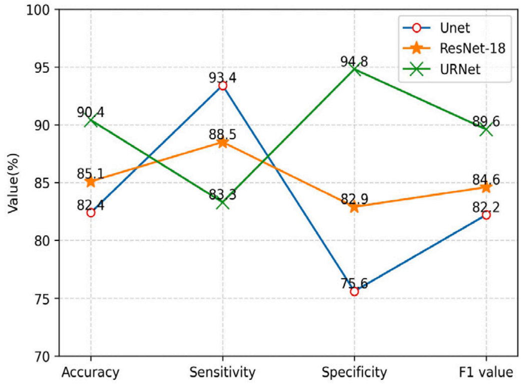


Fig. 7. URNet and baseline models' performance on local dataset.

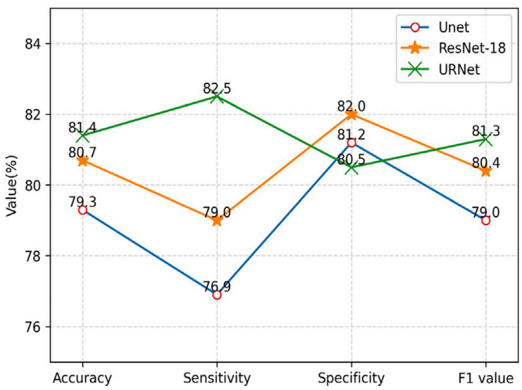


Fig. 8. URNet and baseline models' performance on apnea dataset.

5. Results

To study the influence of the adjacent segment t on the OSA detection of the current period, we conducted experiments on the Apnea-ECG dataset for segment t before and after the current segment. The environment, hyperparameters and evaluation indicators applied in the experiment were the same to ensure the accuracy of the results. Fig. 6 demonstrates that the accuracy is higher when $t = 2$ (five-minute segment).

Ablation experiments were performed based on $t = 2$. Figs. 7 and 8 compare the performance of URNet, Unet and ResNet-18 on both Apnea-ECG and our local dataset, respectively. It is obvious that URNet performs the best. Because the Unet has no specific bright spots, it does not perform well in OSA detection. URNet adds a residual module to the Unet encoder to overcome the degradation and saturation problems. At each layer, the ability of single-lead ECG to detect OSA is improved by extracting more features.

Table 3
The results of the Apnea-ECG database.

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1 (%)
URNet	90.4	83.3	94.8	89.6

Table 4
The offline and online analysis results of the Apnea-ECG database.

Platform	Run time (s)	Power (W)
GTX1080Ti	3.4	253
NVIDIA Jetson Xavier NX	24.2	17

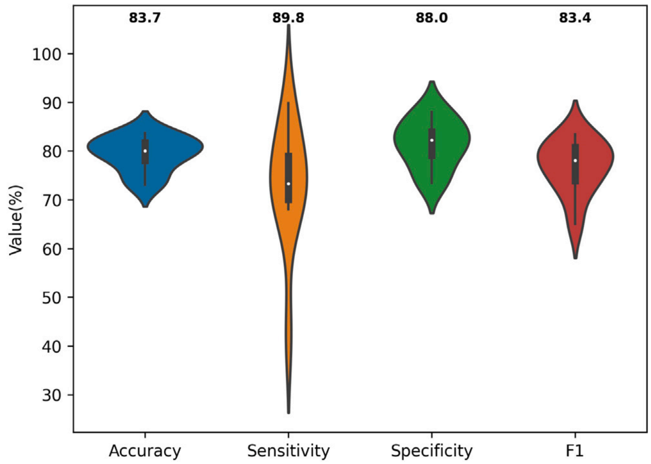


Fig. 9. Ten-fold cross-validation results of local dataset.

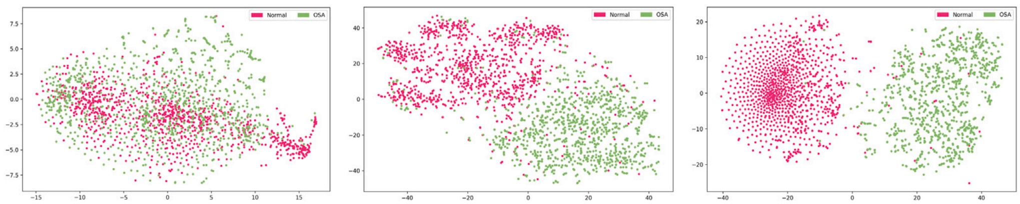
Table 5
The results of different methods using the Apnea-ECG database.

Model	Method	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1 (%)
Machine learning	Sharma et al. [14]	83.3	79.5	88.4	–
	Mendez et al. [27]	85.5	83.9	88.5	–
	Song et al. [28]	86.2	82.6	88.4	–
Deep learning	Wang et al. [16]	87.6	83.1	90.3	–
	Shen et al. [17]*	89.4	**89.8**	89.1	86.4
	Li et al. [18]*	84.7	88.9	82.1	81
	Feng et al. [19]	85.1	86.2	84.4	76.6
	Our method	**90.4**	83.3	**94.8**	**89.6**

The key to OSA detection is the accurate identification of whether OSA occurs within an ECG segment/minute. Two independent datasets were utilized to assess the effectiveness of the URNet. The Apnea-ECG dataset as benchmark data can evaluate the performance of the network. The neural network parameters with the highest accuracy during training were saved as the final network model. As shown in Table 3, the URNet achieves an accuracy of 90.4%, a sensitivity of 83.3%, a specificity of 94.8% and an F1 value of 89.6% on the Apnea-ECG dataset. For the NVIDIA Jetson Xavier NX, we directly load the model with the best accuracy weight to verify the testing set. Therefore, the results of evaluation metrics such as accuracy, specificity, sensitivity and F1 value, were the same as those on PC. Their main differences lie in running time and power consumption, as shown in Table 4. The running time indicates the time required by the algorithm to validate the testing set. We employed Deli power metering socket to calculate the power of GTX1080Ti and NVIDIA Jetson Xavier NX.

Furthermore, we used tenfold cross validation on the local dataset to verify the robustness of our algorithm. The whole dataset is randomly divided into ten parts according to the records. We took the data of each part as the subsequent validation set and used the remaining nine parts for model training. According to Fig. 9, the average of values of accuracy, sensitivity, specificity and F1 are 79.5%, 73.1%, 81.5% and 76.8%, respectively. The summit of each violin represents the maximum value of the corresponding part.

We compared the URNet with previously proposed methods of OSA recognition. Due to differences in dataset and sample sizes, it is not possible to directly compare our results with others. Therefore, all enumeration methods use the Apnea-ECG dataset mentioned in the article. Table 5 lists the overall performance on the test set, including accuracy, sensitivity, specificity, and F1 score. Where



(a) the outputs of Encoder1 (b) the outputs of Encoder4 (c) the outputs of last Decoder

Fig. 10. Visualize features.

∗ represents F1 value calculated by their confusion matrix. Except Mendez which used 50 recordings (25 for training and 25 for testing), others used 70 recordings (35 for training and 35 for testing). It can be seen from Table 5 that the accuracy (90.4%), specificity (94.8%) and F1 (89.6%) of our proposed method reach the highest level. Due to the preprocessing process and the different network input, the corresponding inconsistent data distribution and data segments may affect the sensitivity in our method.

6. Discussion

6.1. Visualize features

We use T-SNE algorithm [29] to visualize features extracted by URNet. The results in Apnea-ECG dataset are shown in Fig. 10. The visualization results indicate that our model has a certain effect on feature discrimination. With the four encoders, the two scattered and indistinguishable categories tend to cluster and the distinction between normal and OSA cases is clear.

6.2. Advantages and limitations

A deep learning model based multiscale feature extraction algorithm is proposed to detect OSA from single-lead ECG. The DOWN_SAMP and CONV blocks are applied to automatically extract effective features from RR intervals. The features obtained by the encoder at each layer are linked to the decoder, so that the network can learn the information lost in the encoder. The residual connection in ResNet introduced in our network not only improves the richness of the features, but also effectively avoids various drawbacks associated with deep layers. Moreover, we also designed a system that can detect obstructive sleep apnea and determine the AHI (severity grade).

However, our work still has some limitations. One is that the study did not differentiate between obstructive, mixed, and central sleep apnea events. This defect will limit its practical application in clinic. Another limitation is that the network only automatically learns the relevant features of OSA from RR intervals, but not from the raw ECG signals.

6.3. Future work

The future research direction is to implement a network model that can directly extract features from raw ECG signals. We believe that through cooperation with local hospital and sleep centers, collecting more and abundant data of sleep apnea events can further improve the performance of our network. Furthermore, health awareness is now shifting from "passive health" to "active health". Collecting ECG signals from wearable devices for simultaneous detection will make daily sleep health monitoring possible in the future. For example, a completely independent wearable ECG device (Shimmer ECGmd) could be applied for integration with the embedded development board for real time OSA detection. The ability to obtain early warning of sleep disorders shows good potential application in daily life.

7. Conclusion

In this study, we design a detection method of OSA with high recognition rate and high efficiency. A multiscale feature extraction algorithm based on embedded deep learning is proposed. The URNet deep learning model has some advantages in automatical detect of OSA events. It can automatically extract important features from signals, while avoiding some traditional machine learning problems and improving the efficiency of feature extraction. The URNet achieves promising classification results with an accuracy of 90.4%, a sensitivity of 83.3%, a specificity of 94.8% and an F1 score of 89.6% on the Apnea-ECG dataset and an average accuracy of 79.5%, an average sensitivity of 73.1%, an average specificity of 81.5% and an F1 score of 76.8% on local dataset. The results on apnea-ECG and local dataset proves that the method has strong robustness. We have also verified it on a embedded device (NVIDIA Jetson Xavier NX). Since the method is based on a single-lead ECG signal, it can be optimally integrated into lightweight device to meat the need for OSA identification outside hospital.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant (62172340 and 61901191), in part by the Natural Science Foundation of Chongqing, China under Grant cstc2021jcyj-msxmX0041, in part by the Fundamental Research Funds for the Central Universities, China under Grant SWU020008, in part by the Young and Middle-aged Senior Medical Talents Studio of Chongqing under grant ZQNYXGDRCGZS2021002. Xinge Yu thanks the Inno HK funding support from the Hong Kong Centre for Cerebro-cardiovascular Health Engineering (COCHE).

References

[1] David P. White, Sleep-related breathing disorder. 2. Pathophysiology of obstructive sleep apnoea, Thorax 50 (7) (1995) 797.
[2] Raphael Heinzer, S Vat, Pedro Marques-Vidal, H Marti-Soler, D Andries, N Tobback, V Mooser, M Preisig, A Malhotra, G Waeber, et al., Prevalence of sleep-disordered breathing in the general population: the HypnoLaus study, Lancet Respir. Med. 3 (4) (2015) 310–318.
[3] Paul E Peppard, Terry Young, Jodi H Barnet, Mari Palta, Erika W Hagen, Khin Mae Hla, Increased prevalence of sleep-disordered breathing in adults, Am. J. Epidemiol. 177 (9) (2013) 1006–1014.
[4] Su Hwan Hwang, Jae Geol Cho, Byung Hun Choi, Hyun Jae Baek, Yu Jin Lee, Do-Un Jeong, Kwang Suk Park, et al., Real-time automatic apneic event detection using nocturnal pulse oximetry, IEEE Trans. Biomed. Eng. 65 (3) (2017) 706–712.
[5] Heenam Yoon, Ji Ho Choi, Hyun Jae Baek, Apneic event estimation only using SpO2 dynamics in sleep apnea patients, in: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), IEEE, 2020, pp. 5335–5338.
[6] Phatchara Temrat, Yuttapong Jiraraksopakun, Apichai Bhatranand, Katinee Wea-asae, Suitable feature selection for OSA classification based on snoring sounds, in: 2018 15th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), IEEE, 2018, pp. 1–4.
[7] Tom Van Steenkiste, Willemijn Groenendaal, Dirk Deschrijver, Tom Dhaene, Automated sleep apnea detection in raw respiratory signals using long short-term memory neural networks, IEEE J. Biomed. Health Inf. 23 (6) (2018) 2354–2364.
[8] T. Douglas Bradley, John S. Floras, Obstructive sleep apnoea and its cardiovascular consequences, Lancet 373 (9657) (2009) 82–93.
[9] Michael Klum, Fabian Leib, Casper Oberschelp, David Martens, Alexandru-Gabriel Pielmus, Timo Tigges, Thomas Penzel, Reinhold Orglmeister, Wearable multimodal stethoscope patch for wireless biosignal acquisition and long-term auscultation, in: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2019, pp. 5781–5785.
[10] Piero Fontana, Neusa Rebeca Adão Martins, Martin Camenzind, Maximilian Boesch, Florent Baty, Otto D. Schoch, Martin H. Brutsche, René M. Rossi, Simon Annaheim, Applicability of a textile ECG-belt for unattended sleep apnoea monitoring in a home setting, Sensors 19 (15) (2019).
[11] Nathaniel F. Watson, Christopher R. Fernandez, Artificial intelligence and sleep: Advancing sleep medicine, Sleep Med. Rev. 59 (2021) 101512.
[12] Chuanhao Zhang, Wenwen Yu, Yamei Li, Hongqiang Sun, Yuan Zhang, Maarten De Vos, Cms2-net: semi-supervised sleep staging for diverse obstructive sleep apnea severity, IEEE J. Biomed. Health Inf. (2022).
[13] Yamei Li, Caijing Peng, Yinkai Zhang, Yuan Zhang, Benny Lo, Adversarial learning for semi-supervised pediatric sleep staging with singleeeg channel., Methods (2022).
[14] Hemant Sharma, K.K. Sharma, An algorithm for sleep apnea detection from single-lead ECG using Hermite basis functions, Comput. Biol. Med. 77 (2016) 116–124.
[15] Daniele Ravì, Charence Wong, Fani Deligianni, Melissa Berthelot, Javier Andreu-Perez, Benny Lo, Guang-Zhong Yang, Deep learning for health informatics, IEEE J. Biomed. Health Inf. 21 (1) (2016) 4–21.
[16] Tao Wang, Changhua Lu, Guohao Shen, Feng Hong, Sleep apnea detection from a single-lead ECG signal with automatic feature-extraction through a modified LeNet-5 convolutional neural network, PeerJ 7 (2019) e7731.
[17] Qi Shen, Hengji Qin, Keming Wei, Guanzheng Liu, Multiscale deep neural network for obstructive sleep apnea detection using RR interval from single-lead ECG signal, IEEE Trans. Instrum. Meas. 70 (2021) 1–13.
[18] Kunyang Li, Weifeng Pan, Yifan Li, Qing Jiang, Guanzheng Liu, A method to detect sleep apnea based on deep neural network and hidden Markov model using single-lead ECG signal, Neurocomputing 294 (2018) 94–101.
[19] Kaicheng Feng, Hengji Qin, Shan Wu, Weifeng Pan, Guanzheng Liu, A sleep apnea detection method based on unsupervised feature learning and single-lead electrocardiogram, IEEE Trans. Instrum. Meas. 70 (2020) 1–12.
[20] Huseyin Nasifoglu, Osman Erogul, Obstructive sleep apnea prediction from electrocardiogram scalograms and spectrograms using convolutional neural networks, Physiol. Meas. 42 (6) (2021) 065010.
[21] Azadeh Yadollahi, Zahra Moussavi, Acoustic obstructive sleep apnea detection, in: 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, 2009, pp. 7110–7113.
[22] Lili Chen, Xi Zhang, Changyue Song, An automatic screening approach for obstructive sleep apnea diagnosis based on single-lead electrocardiogram, IEEE Trans. Autom. Sci. Eng. 12 (1) (2014) 106–115.
[23] Diederik P. Kingma, Jimmy Ba, Adam: A method for stochastic optimization, 2014, arXiv preprint arXiv:1412.6980.
[24] Richard B Berry, Rohit Budhiraja, Daniel J Gottlieb, David Gozal, Conrad Iber, Vishesh K Kapur, Carole L Marcus, Reena Mehra, Sairam Parthasarathy, Stuart F Quan, et al., 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) (2012) 597–619.
[25] Sheng Han, Fan Yang, Hui Jiang, Gang Yang, Na Zhang, Dawei Wang, A smart thermography camera and application in the diagnosis of electrical equipment, IEEE Trans. Instrum. Meas. 70 (2021) 1–8.
[26] Leonel Rosas-Arias, Gibran Benitez-Garcia, Jose Portillo-Portillo, Jesus Olivares-Mercado, Gabriel Sanchez-Perez, Keiji Yanai, FASSD-net: Fast and accurate real-time semantic segmentation for embedded systems, IEEE Trans. Intell. Transp. Syst. (2021) 1–12.
[27] Martin O Mendez, Davide D Ruini, Omar P Villantieri, Matteo Matteucci, Thomas Penzel, Sergio Cerutti, Anna M Bianchi, Detection of sleep apnea from surface ECG based on features extracted by an autoregressive model, in: 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, 2007, pp. 6105–6108.
[28] Changyue Song, Kaibo Liu, Xi Zhang, Lili Chen, Xiaochen Xian, An obstructive sleep apnea detection approach using a discriminative hidden Markov model from ECG signals, IEEE Trans. Biomed. Eng. 63 (7) (2015) 1532–1542.
[29] Laurens Van der Maaten, Geoffrey Hinton, Visualizing data using t-SNE, J. Mach. Learn. Res. 9 (11) (2008).