Automated, effective and efficient sleep-stage monitoring and structure analysis is an essential enabling procedure for healthcare automation. Sleep diagnosis by polysomnography is a golden standard but expensive procedure involving huge effort from patients. There remain challenges for smart devices to precisely identify sleep stage and minimize intrusive effect on sleep progression. Herein, a novel noncontact sleep structure prediction system (NSSPS) using a single radar sensor is presented to analyze sleep structure without any tethered unit. The NSSPS is realized through training a convolutional recurrent neural network and neural conditional random fields using reflected radio frequency (RF) waves acquired by radar antennas. By capturing implicit temporal information in RF signals and transitions of sleep progression, high accuracy of sleep-stage prediction is achieved and characteristics of sleep structure are extracted. The performance of the NSSPS is validated by transfer learning between radar signals with different frequency bands and crossvalidation among different subjects. Moreover, the NSSPS is demonstrated to estimate overnight parameters that are critical for sleep diagnosis. Benefiting from its low cost, convenient setup, and accurate prediction capability of sleep-stage identification, the NSSPS can be widely deployed in “smart” homes and exploited to conduct daily sleep structure analysis.

1. Introduction

Quality sleep is crucial for human health and well-being. Sleep disorders, such as obstructive sleep apnea, have high worldwide prevalence[1] and significantly deteriorate quality of life and cause cardiovascular morbidity and mortality.[2] Most sleep disorders can be managed and prevented with timely diagnostics.[3] Polysomnography (PSG) during sleep is the golden medical standard procedure to diagnose sleep disorders. According to Rechtschaffen and Kales and the American Academy of Sleep Medicine scoring rules, PSG record is segmented into 30 s epochs. Each epoch is scored as one of the sleep stages: awake, rapid eye movement (REM) stage, and three non-REM stages (N1 or N2: light sleep and N3: deep sleep). The sleep stages are correlated with physical repairs, memory consolidation, and immunity enhancement. Sleep structure is generally considered as sleep cycles and stage sequences throughout the night. Monitoring sleep stages and analyzing sleep structure are among the main approaches for diagnosing sleep disorders and assessing the effect of treatment.[4]

To obtain PSG, multiple sensors have to be attached to the human subject and these include electroencephalogram (EEG), electrocardiogram (ECG), chest bands, and peripheral pulse oximetry. These tethered sensors bring intrusive effects on human subjects and potentially disturb their falling asleep. The intrusive effects would further negatively influence their sleep structure and the sleep-quality evaluation results.[5] In addition, the financial cost, time consumption, and strict operation of PSG also limit its usage and availability of long-term sleep monitoring in nonhospital facilities. The diagnostic approach for obstructive sleep apnea should shift from laboratory-based testing to home-based sleep management.[6]
In recent years, smart devices were developed to provide effective, comfortable, and convenient sleep monitoring. Kuo et al.\cite{7} developed a wearable actigraphy sensor, and a high accuracy over 90% for wake-sleep score was achieved compared with manual PSG results. Gu et al.\cite{8} leveraged the built-in acoustic sensors on smartphones to detect a fine-grained sleep stage. The three-stage sleep classification performance was about 64.5%. Walsh et al.\cite{9} evaluated an under-mattress sleep-monitoring system for noncontact sleep/wake discrimination and reported an accuracy of 77.5%. However, these methods need further improvement before they can be adopted for potential practical applications.

The evolution of sensing technologies has led to wireless systems to monitor physiological signals without body contact or tethered connection.\cite{10-12} In these systems, radar sensors transmit a low-power radio frequency (RF) signal and extract human vital signs from the reflective signal. The cardiorespiratory signals captured by radars have been demonstrated feasible to be used for sleep monitoring. In the past decade, researchers have demonstrated great feasibility for using radar sensors to predict sleep stages.\cite{13-16} The radar measurements also provide valuable information for sleep quality evaluation. Compared with traditional PSG methods, such noncontact systems are low cost, user friendly, and convenient to setup, resulting in greater potential for long-term sleep-stage monitoring. In the study by Chazal et al.\cite{17} a novel noncontact biomotion sensor was developed to identify sleep/wake patterns for adults by movement detection and respiration patterns identification. The overall sleep sensitivity was 87.3% and the wake sensitivity was 50.1% against the gold-standard PSG. In the study by Tataraidze et al.\cite{18}, the step-frequency biomedical radar was used for the three-stage (wake, REM, and NREM) classification based on the cycle-based features of respiratory movements, and the average accuracy of 75.1% was achieved. The work in the study by Hong et al.\cite{19} took into consideration of sleep-related signals, including respiration, heartbeat, and body movement. The subspace k-nearest neighbor algorithm outperformed in four-stage (wake, REM, deep sleep, and light sleep) prediction. However, the method using handcrafted features and the classifier was highly dependent on the knowledge of researchers and was sensitive to unknown noise. Therefore, it is difficult to apply these methods to other new sensing techniques and new testing environments.

In sleep analysis with PSG, deep learning approaches have shown high accuracy on sleep staging by eliminating handcrafted features.\cite{19-21} However, sleep analysis based on deep learning and the RF signal still remains in the early development stage. Few studies use and compare different structures of deep learning models carried out with the use of RF signals. Furthermore, existing methods exclusively focus on learning informative features from RF signals and predicting features at the current step. The strong transition structure of sleep states\cite{22} was not thoroughly taken into consideration. Thus, essential dynamics information might be overlooked by the aforementioned learning-based approaches. The goal of this study is to address these knowledge gaps and develop a novel deep learning model with high accuracy and robustness for sleep-stage prediction by noncontact radar sensors.

We present a noncontact sleep structure prediction system (NSSPS) to identify sleep stages. The NSSPS leverages the physiology attributes of respiration, heartbeat, and body movement captured by radar sensors. A neural network (NN) model is used to combine convolutional layers, recurrent layers, and neural conditional random fields (CRF) to train and capture temporal information and describe transition of sleep progression. The robustness and transferability of the proposed method are verified by conducted experiments on diverse RF devices and various subjects. From the proof-of-principle study on human subjects, we have obtained an average accuracy of 75.3% for different sleep stages on 60 GHz samples, and a mean accuracy of 79.2% and Cohen’s kappa of 0.679 are achieved in cross-validation on 6 GHz RF sleep dataset. Compared with methods based on other smart devices and the handcrafted machine learning method, the NSSPS demonstrates a higher detection and estimation accuracy. We further investigate the potential of the NSSPS in calculating sleep parameters and evaluating sleep quality. The results have demonstrated the feasibility to be widely deployed in the “smart” home environment for conducting daily sleep structure analysis. The main contribution of the work lies in a new three-level sleep-monitoring system with millimeter-wave radar sensors and a novel machine learning-enabled sleep transition prediction scheme.

2. Results

2.1. The NSSP System

Figure 1 illustrates the prediction of overnight sleep stages by the NSSPS. The system cycles through sleep stages (i.e., awake, REM, and non-REMs) over a night and each stage has a unique function, that is, REM stage is associated memory consolidation and N3 stage allows human muscles to completely relax. Different functions of stages result in variation of vital signs, as shown in Figure 1a. When sleep progressively deepens in N1, N2, and N3 stages, the breathing and heartbeat slow down and the magnitude of body movements becomes small and less frequent. In the REM stage, heartbeat usually becomes faster and less rhythmical than other stages. The hardware design has a three-level architecture, radar sensors, access point, and cloud, as shown in Figure 1b. The system is developed for real-time sleep monitoring with a radar sensor placed around the bed. The multichannel 60 GHz radar is used to transmit the continuous waves and receive the signals reflected by objects. The human respiration, heartbeat, and limb movement cause variation of the reflected RF signal. Compared with previous RF devices, the radar sensor used in the NSSPS has smaller size, a narrow beam width, and a high signal-to-noise ratio. Moreover, the sensor signals are robust to disturbance generated by other nearby wireless communication devices (as illustrated in Figure S1a,b, Supporting Information).

To obtain the dataset for model training and prediction, the sleep samples are recorded continuously overnight (e.g., 8 h) (refer to the detailed RF sleep samples in “Experimental Section”). Each sample is wirelessly sent to the cloud database. After the subject’s overnight sleep is captured by the noncontact radar system, it is split to construct the training and testing dataset. The supervised model is trained with clinical labels to learn latent representation (Figure 1c). Finally, the trained NN model is used to identify sleep stages for new subjects (Figure 1d). Users can also access results remotely through the graphic interface.
2.2. Sleep-Stage Prediction Method

The sleep-stage prediction task is to annotate each 30 s epoch with a label. Each label takes one of four stages: awake, light sleep (N1 or N2), deep sleep (N3), and REM. Specific definition of the problem is formulated in Supporting Information. As the input of the computational task, RF signals of different frequencies probably have diverse spatial resolutions and reflection coefficients. To eliminate the specific differences of sensors and vital signs, components in the RF signals are extracted and normalized (refer to the detailed description of RF signal preprocessing in Supporting Information).

As shown in Figure 2, we introduce an NN model to implement sleep-stage prediction. It is composed of a convolutional-recurrent neural network (CRNN) to learn the latent representation, a multichannel voting mechanism to estimate the reliability of channels, and a CRF layer to model the transition process of the sleep state. The architecture of convolutional neural network (CNN)–recurrent neural network (RNN) is designed as an encoder to extract time-invariant features from each 30 s epoch and learn sequences of epoch series. A CNN is used to capture information of different timescales and its structure is illustrated in Figure 2a. Inspired by signal processing technology, we use two 1D convolution kernels of different sizes in the first layer to extract temporal information from different-frequency components. The CNN consists of 16 1D convolution layers and each 1D convolution layer is a sequence of filters, batch normalization, and rectified linear unit activation.

As shown in Figure 2b, an RNN structure, simple recurrent units (SRU), is used to capture the dynamics of features and learn feature transition rules. For instance, the light sleep stage lasts for about half hour per cycle and then the body usually falls into a deep sleep. In this case, RF features of vital signs under light and deep sleep resemble each other, while they can be distinguished from awake or REM by the CNN. Then RNN is capable of learning to remember the sleep history and deliver a summary to the current cell.

A CRF layer is treated as a global predictor at the end of the NN model. Resulting from the multipath effect and noise observed by different RF channels, extracted features probably vary from channel to channel. As shown in Figure 2c, a voting mechanism is used to assess weights for each channel representing its relative significance. The CNN–RNN encoder extracts the best features from the RF signals and the sleep-stage prediction depends on each time step. However, it is known that the sleep-stage transitions have a strong dependency structure. For example, a deep sleep stage cannot be reached without going through a light sleep stage. This transition structure for accurate sleep staging is defined in Figure 2d. With the CRF model as shown in Figure 2e and the joint conditional probability (Equation (11) in the Supporting Information), we reach a globally optimal decision of sleep stages. Details on the NN model are provided in Supporting Information.

2.3. Sleep-Stage Prediction Results

As for real applications, the overnight sleep data normally come from new individuals, new environments, and even new devices. Variations exist in the physiological features under sleep stages between different subjects, ages, health status, etc. To test the classification capability of the NSSPS, we perform transfer learning validation on different sensors and human subjects. Prediction accuracy and Cohen’s kappa κ are computed to evaluate the performance (refer to evaluation metrics in “Experimental Section”). Overall, the proposed method finally performs with an overall accuracy of 79.2% and κ value of 0.679. It has demonstrated capability to learn knowledge of sleep structure from training set and extract information from RF signals. Figure 3a shows an example of sleep structure labeled by experts.
as ground truth and predicted by the model. Accuracy of the example is around the mean value. Figure 3b shows the overall confusion matrix. The prediction accuracies for awake, REM, light sleep, and deep sleep stages are 66%, 80%, 83%, and 76%, respectively. The identification of light sleep obtains the highest accuracy. The worst performance is noted for awake stage. The awake stage is mostly mistakenly identified as REM or light sleep stages. Other misclassifications are mainly between light sleep and deep sleep stages. These two stages are both non-REM and have similar physiological features.

To verify the representation capability of the NN model, we performed testing on new devices. However, the trained model without fine tuning failed on the new device. Meanwhile, training merely with smaller amount of data and random model initialization had a low speed and accuracy bottleneck. Therefore, we pretrained the CRNN encoder with 6 GHz RF dataset and performed a transfer learning method to fine tune the end-to-end NN model with 60 GHz RF signal. Performance on each subject is evaluated using the model fine tuned with data from other subjects. Figure 3c shows the accuracy and $\kappa$ values on each subject.

Figure 2. The NN model to implement sleep-stage prediction. a) A CNN to extract temporal information in 30 s epochs. b) An RNN module to learn sequential knowledge. c) A multichannel voting mechanism to estimate logits of classes. d,e) A CRF layer to model the transition process of the sleep state.
and new radar sensor. The mean values of accuracy and $\kappa$ value are 74.3% and 0.594, respectively. The transfer learning result demonstrates that the model is capable of being used in different devices with a few training steps.

The robustness of NSSPS is also evaluated using crossvalidation on MIT dataset. Figure 3d illustrates the $\kappa$ value on different subjects achieved by the CRNN and CRF model. It has a standard deviation of 3.1%. The maximum and minimum accuracy rates are 83.2% and 72.8%, respectively. The $\kappa$ value has a standard deviation of 0.053. The maximum and minimum $\kappa$ values are 0.736 and 0.552, respectively. The crossvalidation and transfer learning results have shown that the proposed method has a high potential for real applications.

To explain the extraction mechanism of the encoder, Figure 4a,b shows the visualization of the input saliency and encoder output, respectively. These plots help understand how the NN model works. Simonany et al.\cite{25} proposed an approach to take the gradient of the classification scores with respect to the input as the saliency map, which illustrates where the model is “looking” at. As shown in Figure 4a, two components (respiration and heartbeat) in RF signals are plotted with a background of the corresponding saliency map. Darker lines in the saliency

![Figure 3](image-url)

**Figure 3.** Sleep-stage prediction results. a) An example of sleep-stage prediction. b) Overall confusion matrix of results. c) Prediction result of the new frequency-band device and new subjects. d) Cohen’s kappa of crossvalidation on different subjects in MIT dataset.

![Figure 4](image-url)

**Figure 4.** a) Saliency map where darker lines represent higher concentration of the network. b) Visualization of features extracted by the CRNN. c) Probability transition matrix of the previous sleep stage and next stage where zero values represent prohibition of the corresponding stage transitions.
map denote higher attention of the encoder. It can be seen that the model concentrates on peaks and valleys of the respiration and heartbeat signals. With small-sized and large-sized filters in the first convolution layer of the encoder, cycles of vital signs and the variation of physiology are effectively captured. We employ t-SNE embedding to visualize the outputs from the CNN and RNN separately, as shown in Figure S2, Supporting Information. The CNN extracts the temporal features at each 30 s epoch and has a primary discrimination of awake and REM from non-REM stages. However, vital signs under light sleep and deep sleep resemble each other and the CNN fails in making decisions between these two stages (Figure S2a, Supporting Information). The RNN output (Figure S2b, Supporting Information) has obviously clearer boundaries of classes than the CNN features. The RNN model learns dynamics from the sleep sequences and enables to further determine the depth of sleep. Moreover, concatenation of the CNN output enhances the latent learning of the awake stage (red points in Figure 4b and S2, Supporting Information). We trained a similar model without concatenation of these two types of features and the overall accuracy rate decreases by 2.9%. This implies that the concatenation improves the training progress and enhances the model classification ability. As for the predictor of the CRF layer, the stage transition matrix learnt from the training data shows the probability of adjacent stage pair, as illustrated in Figure 4c. The zero values in the matrix prohibit the corresponding stage transitions (e.g., from deep sleep to REM) in the prediction output. The CRF model comprehensively utilizes the features of channels, joint transition probability, and makes a globally optimal decision. Compared with the CRNN model, the sleep-stage transition model improves an accuracy of 2.6% and a 𝜅 value of 0.036. A detailed comparison over different subjects is shown in Figure 3d. The visualization results of the NN model confirm the method to be explicable and credible.

### 2.4. Sleep Structure Analysis

Sleep structure parameters are relevant to sleep quality. Accurate estimation of these parameters is useful for both the patient and medical practitioner to monitor progress on clinical therapy. The following six sleep parameters are chosen to verify NSSPS’s feasibility in sleep structure analysis: Total sleep time is defined as the total time of nonwake stages. Sleep efficiency is defined as the proportion of non-awake stage during a whole night of sleep to assess sleep quality. Sleep latency and REM latency are defined as the elapsed time from the start of on-bed to the first 30 s epoch scored as sleep and REM, respectively. Deep sleep and REM proportions are defined as overall percentages of the deep sleep stage and REM stage during a whole night, respectively.

The aforementioned parameters are calculated on all the overnight sleep data sampled from the radar sensors and the PSG method. Figure 5 shows the parameters’ distribution of health subjects and the estimation-truth comparison between the RF and PSG methods. The dotted lines in each figure denote the upper and lower quartile of the signed estimation errors. The absolute estimation error and mean accuracy on different parameters are illustrated in Table S1 and S2 in Supporting Information. Specifically, the estimates of the total sleep time and sleep efficiency achieve the highest accuracy. Although it is obviously less accurate than other parameters, the sleep latency estimation results from a short elapsed time before the subject falls asleep. The corresponding absolute error is notably small on average and verifies feasibility in identifying the awake–sleep transition. Accuracy rates of other estimated parameters are all higher than 70%. The results demonstrate the capability of the NSSPS in estimating sleep parameters for sleep analysis and its great potential in clinical sleep diagnosis.

### 3. Discussion

In this proof-of-concept study, the capability of the NSSPS for sleep-stage prediction and sleep structure analysis has been demonstrated on radar sensors with different operating frequency and subjects. With the proposed NN model, high sleep-correlative vital signs in RF signals, and a considerable amount of data, the NSSPS outperformed other state-of-the-art sleep-stage prediction methods. The preliminary result also demonstrated the robustness on multiple human subjects and different sensor devices. We further verified that the NSSPS can analyze the overnight sleep structure and has a potential in clinical applications.

#### 3.1. Result Comparison

We compared our results with other reported methods using different sensors, as shown in Table S3 in Supporting Information. Compared with the sensing devices such as smartphones, actigraphy, and nasal airflow, RF signals contain additional information of vital signs, and the NSSPS improves the sleep structure estimation accuracy significantly. Meanwhile the proposed system is more convenient to use than wearable sensors and the results are more reliable than smartphones that collect acoustic signals. Moreover, the stage prediction accuracy is comparable with the EEG-based method that used electrodes attached to the scalp. The EEG devices require stable contact with scalp and may cause discomfort. The advantages of the NSSPS lie in convenience, reliability, and contactless setup. Table S4 in Supporting Information presents the comparison of accuracy on two different RF sensors. Benefiting from the high quality of millimeter wave, the proposed system distinguishes the awake stage accurately. We also compared the proposed NN model with other prediction models using RF signals, as shown in Table S5 Supporting Information. A comparison with the handcrafted features and classifier in the previous studies demonstrates that the proposed machine learning-enabled method preferably extracts the features and outperforms in sleep-stage prediction. For example, the comparison with the classic CNN–long short-term memory (LSTM) network, as shown in Figure 3d, implies that the utilization of the CRF achieves a globally optimal decision and improves the accuracy by 2.6%. This improvement benefits from fusion of multichannel RF signals to strengthen the training of NN and enhance robustness to environment noises. The prediction accuracy in this work is also comparable with the conditional adversarial architecture in the study by Zhao et al. in which a universal training regime discards individual-specific information.
One advantage of the proposed NN model also lies in high-computing speed in the overnight sleep-stage prediction. The conventional RNNs suffer heavy computation load when they are applied to long sequences. With the highly parallel SRU cells, the time consumption in the proposed NN model during training and testing progress is decreased by about 50% than that of LSTM or gated recurrent unit (GRU). Furthermore, we implemented model training on 60 GHz dataset from the pretrained model on the 6 GHz dataset and random initialization, respectively. The comparison of the learning process is illustrated in Figure S3 in Supporting Information. The pretrained model and the transfer learning approach remarkably improve the prediction performance with the two-step training algorithm. It demonstrates the common knowledge in different RF signals and transferability of the NN model.

As for sleep structure analysis, the proposed method also outperforms the sleep efficiency estimation reported in the study by Aggarwal et al.[20] With the demonstrated higher accuracy of sleep-stage prediction and structure parameter estimation, the NSSPS has higher capability of exporting sleep cycle measurement and sleep event detection (e.g., sleep arousal). By providing more effective clinical information, it has great potential for sleep quality assessment with simplified testing procedures.

### 3.2. Potential Applications

The NSSPS has many attractive advantages, such as miniaturization, low cost, and a convenient setup. The results in this study show the potential of the NSSPS as a fully automated, noncontact, and nonprivacy-invasive modality for sleep monitoring. The system is adaptable for various scenes and particularly attractive to be widely deployed in "smart" home applications. The users can be aware of their own sleep health in the long term and get rid of the confinement of limited medical resources. In this way, detection of evolution during the course of therapy or timely discovery of deterioration is achievable.

Furthermore, the accurate evaluation of sleep disorders is essential for effective therapy. The accurate sleep structure analysis of the proposed system can be combined with other RF-based in-door monitoring algorithms (e.g., sleep apnea detection, heartbeat analysis, and indoor tracking). With long-term sleep disorders’ evaluation conducted by the NSSPS, doctors are
capable of comprehensively determining the impact of disorders on patients and grade severity more precisely. Thus, more effective therapy can be formulated.

3.3. Limitations and Future Work

It is important to note that the prediction method in this work relies on learning latent representation in RF signals. Some limitations remain in actual experiments in this study, mainly due to labeling data. PSG testing has various degrees of impact on subjects (e.g., difficulty in falling asleep and easy to awake) and might change their sleep habits. In a few epochs when the PSG leads are in poor contact, the corresponding sleep stages are labeled dependent on the doctors’ experience. The limitations might bring individual specificity and accidental error, which restrict the training of the NN in experiments. In addition, the unbalanced class distribution (e.g., less deep sleep) limits prediction performance on sleep stages that have obviously fewer occurrences in training data.

For future work, we intend to implement additional clinical trials of patients. By tracking the course of therapy and developing stronger deep learning engine (e.g., transfer learning), we could improve the system performance of evaluating patients’ evolution. Furthermore, we plan to develop recognition algorithm of sleep apnea with NSSPS and comprehensively analyze sleep disorders of patients. Eventually the presented system should be developed to satisfy long-term demands for convenient in-home sleep monitoring and relieve medical pressure of clinical diagnosis.

4. Experimental Section

RF Sleep Samples: The overnight RF sleep samples were captured from eight healthy adults from Zhejiang Hospital and 25 subjects from datasets provided by Massachusetts Institute of Technology (MIT).[27] Each subject slept with a wireless sensor fixed around and a PSG-based (SOMNOScreen plus, SOMNOMedics GmbH, Germany) sleep monitor to provide the ground truth (refer to the experimental scene, as shown in Figure S4, Supporting Information). Referring to the electronic records of PSG (e.g., EEG, ECG, nasal airflow, and oximetry), experts labeled every 30 s of sleep with one of four sleep stages: awake, REM, light sleep (N1 or N2), and deep sleep (N3). The PSG-based sleep monitor had human-level comparable accuracy.[28] The study was approved by the Medical Ethics Committee of Zhejiang Hospital (Approval Letter NO: AF/SC-06/04.2).

Training of the Prediction Model: We evaluated the performance of the proposed method by performing transfer learning on different devices and leave-one-out-cross-validation (LOOVC) on different subjects. For LOOVC, the model was tested on one subject while training on the data from other subjects. To further identify the performance, we tested the trained model on RF signals sampled by devices of the new frequency band (i.e., device of 60 GHz, which was different from around 6 GHz in the MIT data set) and sampled from new subjects.

To train the model, we implemented a two-step algorithm. The first step was supervised pretraining on a CRNN encoder, whose parameters were denoted as \( \theta_{\text{CRNN}} \). As the class distribution of sleep stages was unbalanced (i.e., REM sleep was less than 10% of annotations), the CRNN encoder was trained with a minibatch gradient-based optimizer called stochastic gradient descent and a cost-sensitive loss function \( \mathcal{L}(\theta_{\text{CRNN}}) \). The second step was to perform an end-to-end training on the parameter \( \theta \) of the entire model with a global loss function \( \mathcal{L}(\theta) \). Specifically, the parameters of the encoder net were initialized by the \( \theta_{\text{CRNN}} \) values that were obtained in the first step. The detailed training process are provided in Supporting Information and Algorithm S1 shows the flowchart.

In addition, we used two techniques to prevent overfitting. One technique was \( L2 \) weight decay in loss function to prevent large values of the parameters (i.e., exploding gradients). The weight decay parameter \( \lambda \) was applied on the CRNN and voting net, and it was set to 0.001. The other technique was dropout, which randomly omitted a fraction of the hidden units.[29] Dropout layers with the probability of 0.2 were added to each SRU layer, while dropout layer with the probability of 0.3 was used to randomly omit the CRNN features. It is important to note that these dropout layers work only for training, and they are inactive during testing.

The implementation used Pytorch in experiments with different model structures and hyperparameters: kernel sizes \( \{5,7,9,15,31\} \) in the first layer, residual module layers \( \{2,3,4\} \), RNN layers \( \{1,2,3\} \), and RNN hidden size \( \{128,256,512\} \). We chose learning rate 0.005 from \( \{0.01,0.005,0.0001\} \) and regularization coefficients 0.001 from \( \{0.005,0.001,0.0001\} \).

Evaluation Metrics: Commonly for the classification task, accuracy is defined as the correct percentage of sleep-stage prediction agreeing with the ground truth. Moreover, we used specific metrics in automated sleep staging, namely, Cohen’s kappa, to compare our model’s prediction with ground truth from PSG. Cohen’s kappa coefficient (\( \kappa \)) was commonly used to measure the inter-rater reliability.[30] It is robust as Cohen’s kappa takes into consideration blind luck in prediction. Coefficient \( \kappa \) calculated the concordance between the sleep-stage prediction and the PSG-based ground truth and had values between 0 and 1. Scores \( \kappa > 0.4 \), \( \kappa > 0.6 \), and \( \kappa > 0.8 \) were considered to be moderate, substantial, and in almost perfect agreement, respectively.[31]

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

This work was supported in part by the National Science Foundation of China, awards 52175033, 51775485, and U21A20120, and the Zhejiang Provincial Natural Science Foundation under award LZ20E050002. The authors are grateful to Dr. Mingmin Zhao for allowing to use sleep dataset.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords

deep learning, healthcare systems, noncontact radar sensors, radars, sleep monitoring

Received: November 3, 2021

Revised: December 24, 2021

Published online: February 4, 2022

[1] A. Benjafield, N. Ayas, P. Eastwood, R. Heinzer, M. Ip, M. Morrell, C. Nunez, S. Patel, T. Penzel, J.-L. Pépin, P. Peppard, S. Sinha, S. Tufik, K. Valentine, A. Malhotra, Lancet Respir. Med. 2019, 7, 687.
