A lightweight automatic sleep staging method for children using single-channel EEG based on edge artificial intelligence

Liqiang Zhu1 · Changming Wang2,3 · Zhihui He4 · Yuan Zhang1

Received: 27 July 2021 / Revised: 8 November 2021 / Accepted: 26 November 2021 / Published online: 30 December 2021
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract
With the development of telemedicine and edge computing, edge artificial intelligence (AI) will become a new development trend for smart medicine. On the other hand, nearly one-third of children suffer from sleep disorders. However, all existing sleep staging methods are for adults. Therefore, we adapted edge AI to develop a lightweight automatic sleep staging method for children using single-channel EEG. The trained sleep staging model will be deployed to edge smart devices so that the sleep staging can be implemented on edge devices which will greatly save network resources and improving the performance and privacy of sleep staging application. Then the results and hypnogram will be uploaded to the cloud server for further analysis by the physicians to get sleep disease diagnosis reports and treatment opinions. We utilized 1D convolutional neural networks (1D-CNN) and long short term memory (LSTM) to build our sleep staging model, named CSleep-Net. We tested the model on our childrens sleep (CS) dataset and sleep-EDFX dataset. For the CS dataset, we experimented with F4-M1 channel EEG using four different loss functions, and the logcosh performed best with overall accuracy of 83.06% and F1-score of 76.50%. We used Fpz-Cz and Pz-Oz channel EEG to train our model in Sleep-EDFX dataset, and achieved an accuracy of 86.41% without manual feature extraction. The experimental results show that our method has great potential. It not only plays an important role in sleep-related research, but also can be widely used in the classification of other time sequences physiological signals.

Keywords Sleep staging · Edge AI · Deep learning · LSTM · EEG

Liqiang Zhu and Changming Wang contributed equally to this work.

This article belongs to the Topical Collection: Special Issue on Resource Management at the Edge for Future Web, Mobile and IoT Applications
Guest Editors: Qiang He, Fang Dong, Chenshu Wu, and Yun Yang

✉ Yuan Zhang
yuanzhang@swu.edu.cn

Extended author information available on the last page of the article
1 Introduction

At present, most of the computing tasks of the remote sleep monitoring (such as sleep staging) are deployed on platforms with large-scale computing resources such as computing centers, which largely limit the convenience to people. With the advancement of Internet of Things (IoT), the number of networking devices has increased dramatically, and the total amount of data generated has increased accordingly. Data transmission from the end to the cloud will bring huge bandwidth pressure and energy consumption, which makes the traditional centralized processing unbearable, thus giving birth to the edge computing and gradually developing to the edge AI [15, 25]. Sleep staging is a basic work of sleep research, which requires the use of sleep records throughout the night. Automatic sleep staging based on edge artificial intelligence can solve the problem of traditional remote sleep monitoring that sleep staging on a cloud server consumes a lot of network resources.

Sleep-induced diseases such as insomnia, drowsiness, obstructive sleep apnea (OSA) and other sleep disorders are becoming more and more common and have become a major medical challenge. For children, high-quality sleep helps children’s intellectual development and is closely related to children’s cognitive function, learning and attention. If school-age children are not able to get enough and good sleep, it will affect their mental development and cause emotional, behavioral, and attention problems. However, nearly one-third of children suffer from sleep disorders [11].

Polysomnography (PSG) recordings is used to diagnose sleep-related diseases, which include electroencephalogram (EEG), electrooculogram (EOG), electrocardiogram (ECG), electromyogram (EMG), breathing exercises (chest and abdominal), oral and nasal airflow, body movement, blood oxygen saturation (SaO2) and other physiological parameters. Sleep stage scoring is to divide the physiological parameters in the polysomnography chart into 30 s continuous epochs according to the time axis, and divide these epochs into different sleep stages according to the American Academy of Sleep Medicine (AASM) rules [4]. Sleep stage scoring can be performed using single-channel EEG or multiple physiological parameters. The hypnogram obtained from the results of sleep staging can intuitively reflect the sleep of subjects throughout the night, and is used to evaluate sleep quality and sleep-related problems [6].

The AASM manual was first published in 2007 [4]. In the AASM rules, all sleep recordings are divided into 5 stages. They include Wake (W), Rapid Eye Movement (REM) and Non-Rapid Eye Movement (NREM), where the NREM includes N1 (transition stage), N2 (light sleep) and N3 (deep sleep). Age may be the most critical factor in differentiating the sleep pattern between children and adults, due to the EEG variation reflected by PSG monitoring [10]. The AASM rules also include sleep stage scoring methods for children.

Technicians need to spend a lot of time and efforts to monitor the changes of different physiological signals in the PSG for sleep stage scoring. In addition, the quality of sleep stage scoring depends on the experience and fatigue of technicians, and the agreement between the technicians is usually less than 90% [28]. In addition, the existing automatic sleep stage classification methods are for adults by default. Therefore, it is necessary to develop an automatic sleep staging method for children.

To perform automatic sleep staging on a cloud server, it is necessary to upload the collected sleep recordings of the whole night (usually hundreds of MB or even more than 1 GB size) to the cloud, which takes a long time to respond, puts a lot of pressure on network bandwidth and may leak user privacy. Edge AI trains and deploys deep learning models at the edge of the network closer to users and data sources, thereby improving the
performance and privacy of AI applications [25]. Therefore, we develop a lightweight automatic sleep staging method for children using single-channel EEG based on edge AI. The sleep staging performed on edge devices, and the results and hypnogram will be uploaded to the cloud server for physicians to further analyze. Users will get analysis reports and some useful suggestions.

COVID-19 has made a huge impact on people’s work, study and life [2]. The rapid development of wearable computing technologies has led to an increased involvement of wearable devices in the daily lives of people, and promoted the development of telemedicine [26, 33]. Telemedicine can effectively reduce the risk of being infected with new coronary pneumonia and other diseases. Our proposed approach can perform sleep monitoring at home instead of in the hospital. It both provides better protection for children and their families, and eases the strain on medical resources.

We conducted experiments on our CS dataset and public dataset sleep-EDFX and achieved satisfactory results. We summarize our contributions as follows:

- We design our automatic sleep staging model for children that utilizes 1D-CNN and LSTM. The 1D-CNN can be trained to learn and extract features from raw single-channel EEG, while the LSTM can be trained to learn temporal information such as sleep stages transition rules.
- We develop a lightweight automatic sleep staging method for children based on edge intelligence. The sleep staging process is carried out on the intelligence terminals, thereby improving the performance and privacy of sleep staging application.

The remaining parts of this paper are organized as follows: First, we review the related works of edge intelligence and automatic sleep staging in Section 2. Then, the automatic sleep staging method for children based on edge AI is proposed in Section 3. Section 4 describes datasets, data processing, the experiments and analysis results. Finally, Section 5 presents the conclusion and future work.

2 Related works

2.1 Applications of edge AI

At present, edge AI has been applied in many fields such as autonomous driving, emotion recognition, object detection and object recognition. Rincon et al. [23] adapted edge AI technologies for the classification of emotions for humans. Uddin et al. [30] proposed a novel emotion recognition method using audio speech and machine learning (ML). And they pushed the processing of data and making decisions towards where data sources. Liu et al. [21] proposed a relay-based data distribution scheme for cell-assisted autonomous driving based on edge AI. Khandewale et al. [20] applied edge AI and neural compute to develop an object detection system for visually impaired people, and the system will support them for person recognition in future. Huang et al. [18] adopted the edge AI philosophy distributing supervised machine learning, data processing and decision-making tasks among the physical layer, edge layer and cloud layer. They proposed a Digital Twin-driven anomaly detection framework to enable industrial systems to perform anomaly detection in real-time. Zhen et al. [35] proposed a application framework for edge AI, called CareEdge and built an ECG-based heartbeat detection system through it. Their experimental result
shows that CareEdge is light and fast than than similar frameworks. Huang et al. [17] proposed an edge AI framework for building intelligence IoT applications which named WuKong. And they implemented an online recognition case to verify the performance of this edge AI framework. El-Shal et al. [33] applied edge AI to detect license plates from real-time video streams.

The emergence and development of edge AI has greatly eased the problem of network resources and improved the performance and privacy of AI applications. Edge AI has been applied to many fields with broad application prospects and has become a development trend. Applying edge AI to children's sleep staging is of great significance for children's daily life sleep health monitoring.

2.2 Methods of automatic sleep staging

In the past few decades, scholars have proposed some automatic sleep staging methods based on machine learning [1, 3, 7, 12–14, 24]. Agarwal and Gotman [1] applied Maximum Overlap Wavelet Transform and Shift Invariant Transform to extract features in the time and frequency domains, and then applied Support Vector Machine (SVM) for sleep stage classification. They obtained accuracy of 80.6% in ISRUC-Sleep dataset. Estrada et al. [7] proposed three different schemes to extract the characteristics of EEG signals: relative spectral band energy, harmonic parameters and Itakura distance. See et al. [24] applied sample entropy and the power spectrum of the harmonic parameters of the infinite impulse response filter and wavelet transform to extract features from the EEG data obtained from the Physionet database, and applied SVM for sleep stage classification. The preliminary result achieved accuracy of 96.2%. Hassan et al. [12–14] used Tunable-Q factor Wavelet Transform (TQWT) to decompose EEG signals to extract various spectral features, and applied adaptive boosting, random forest and bootstrap aggregating for sleep stage classification on the Sleep-EDF dataset. They obtained overall accuracies of 91.94%. Alickovic et al. [3] used multi-scale principal component analysis to denoise Pz-Oz channel EEG signal, and used the discrete wavelet transform (DWT) to extract the most informative feature. The extracted features were input into the integrated classifier called rotational SVM (RotSVM). This classifier combines the advantages of principal component analysis and SVM to improve the classification performance. For the five-stage sleep stage classification, they obtained the sensitivity of 84.46% and accuracy of 91.1%, and the Cohen’s kappa coefficient was 0.88.

In recent years, deep learning has used multi-layer linear and non-linear processing units to learn rich feature representations from raw input data, and has achieved impressive results in various fields such as computer vision and natural language processing. Similarly, deep learning algorithms have also been applied to sleep staging [8, 16, 27, 29, 31, 32, 34]. Hsu et al. [16] extracted energy features from the Fpz-Cz channel EEG signal, and proposed a recursive neural classifier based on energy features for sleep staging. They obtained the accuracy of 87.2% in the Sleep-EDF dataset. Zhang et al. [32] combined complex-valued anti-propagation and Fisher criterion, and proposed a new model called fast discriminative complex-valued convolutional neural network to learn discriminative features, and overcome the negative effects of unbalanced data sets. The total accuracy and Kappa coefficient of this method were 92% and 0.84, respectively. Supratak et al. [29] proposed a deep learning model DeepSleepNet based on the original single-channel EEG, which used convolutional neural networks to extract time-invariant features, and used bidirectional memory to automatically learn the transition rules between sleep stages from the
EEG cycle. The overall accuracy and macro F1 score of the model on both datasets were: MASS: 86.2%-81.7, Sleep-EDF: 82.0%-76.9. Sors et al. [27] used 14-layer CNN to perform supervised learning of 5-stage sleep stage classification based on single-channel EEG. They took the previous epoch and the next epoch of the target epoch together with the target epoch as input. Using data from Sleep Heart Health Study (SHHS) to train and evaluate the model, the model obtained an accuracy of 0.87 and a Kappa coefficient of 0.81. Fraiwan et al. [8] researched the application of LSTM learning system in the automatic sleep stages scoring. The LSTM network architecture was built using Uni-directional and Bi-directional structures to utilize both the forward and backward chains of data sequences. According to AASM standard, the overall accuracy was 91.92%, and the Cohens kappa value was 77.73%. Zhang et al. [31] developed a new unsupervised competitive convolutional neural network (C-CNN), which overcomed the difficulty of obtaining label data. The convolution operator was used to extract the characteristics of the EEG signal, and the competition layer iteratively adjusted the weight vector of the winning neuron according to the competition learning rules. In this way, the learned weight vector can reflect the distribution of input samples. The classification performance of this model on UCD and Sleep-EDF datasets were 77.2% and 83.4%, respectively. We propose a novel hybrid manifold-deep convolutional neural network with hyperbolic attention. To overcome the shortage of labeled data, we update the semi-supervised training scheme as an optimal solution. In order to extract the latent feature representation, we introduce the manifold learning module and the hyperbolic module to extract more discriminative information. Our method yields the overall accuracy of 89% in Sleep-EDFX dataset [34]. The proposed model demonstrates powerful ability in extracting feature representation and achieves promising results by using semi-supervised training scheme. These methods verify the practicability and effectiveness of deep learning for automatic sleep staging.

2.3 Summary

By summarizing the above work, we found that edge AI currently has no relevant applications in the field of sleep staging. The traditional machine learning methods rely on manual feature extraction, and the quality of feature extraction directly affects the results. For sleep staging, researchers manually extract corresponding features based on the characteristics of EEG signals, and the selection of features is affected by the dataset, resulting in poor generalization. Deep learning is an end-to-end learning process that integrates feature extraction and classification into an algorithm, which can overcome the limitations of manual feature extraction. For sleep staging, although the accuracy of current deep learning methods is generally not as good as machine learning, they can independently learn EEG characteristics, have better generalization capabilities, and have more development potential.

In addition, the existing research results aim at adult sleep stage scoring, and there is a lack of research on automatic sleep staging for children. However, children and adults have different EEG characteristics so that these methods are not necessarily suitable for children. Therefore, it is necessary to develop a sleep staging method is more suitable for children [10]. At the same time, PSG and sleep EEG signal of children are few, and childrens sleep monitoring is difficult. In the process of sleep monitoring, children are more sensitive to the monitoring equipment, and children will feel discomfort due to the monitoring equipment, so noises such as artifacts appear and the equipment may even fall. Therefore, the study of sleep stage classification for children in terms of edge AI environment is worthy of long-term research and exploration.
3 Methods

In this section, we firstly present an overview of our CSleepNet model. Then, we provide description of convolutional block (CB) and LSTM in detail. Next, we introduce the loss function and optimization. Finally, we present automatic sleep staging based on edge AI.

3.1 CSleepNet

The architecture of our CSleepNet model is shown in Figure 1. The 30 s EEG epoch is input sequence. The CB is trained to extract time-invariant features from each epoch. The output of the CB is a high-level feature map, which is then used as the input of the LSTM. The LSTM is trained to encode the temporal information. There are three fully-connected layers, with 1024, 256, and 64 neurons after LSTM layer. Then, the soft-max layer outputs the predict results of sleep staging. Comparing the predicted value with the target value and calculating, we can get the performance of sleep staging.

All convolutional layers and fully connected layers use a Rectified Linear Units (ReLU: \( f(x) = \max(0, x) \)) as the activation function, which is not easy to reach its saturation state to avoid the dispersion of the gradient, reducing the amount of calculation and alleviate the occurrence of overfitting. In addition, in order to reduce overfitting risks, the dropout layer is used before the LSTM layer, the third fully connected layer and the soft-max layer, respectively. The implementation details of the CSleepNet structure are shown in Table 1.

3.2 Convolutional block

The CB is composed of 7 convolution layers, 3 max-pooling layers and 3 batch normalization layers, and used to extract features of raw single-channel EEG. We describe implementation details of the CB by referring to Figure 2. The input data of the CB is 30 s EEG epoch, which is a \( 1 \times (30 \times f) \) one-dimensional time sequence, where \( f \) represents EEG frequency. Therefore we applied 1D convolutional kernel to replace 2D convolutional kernel, and adjusted the network structure according to our data characteristics. The output of the CB is a highly abstract feature map, which is then used as the input of the LSTM. The one-dimensional convolution operation is defined as:

![Figure 1](image-url)
where $y_i^l$ is the $i$-th feature map of the output feature map set on the layer $l$. $\omega_n^l$ and $b_l^i$ represent the weight vector and bias unit of the convolution kernel of the layer $l$, respectively. $d$ is the size of the convolution kernel. $N$ is the length of the input feature vector $y_{l-1}^i$.

$$y_i^l = \sigma_{g_{p'}} \left( \sum_{n=1}^{d} \omega_n^l \cdot y_{n+1}^{l-1} + b_l^i \right), i \in (1, N - d + 1)$$  \hspace{1cm} (1)
represents a $p^l$-strided sub-sampling operator, and $\sigma$ is the activation function of the convolutional layer $l$.

The convolutional kernel of the first four convolutional layers have 17 size, and the convolutional kernel of other convolutional layers have 15 size. Although using multiple convolutional layers with smaller convolution kernels is less computationally expensive, too small a convolution kernel does not seem to be able to learn the features in the EEG signal well, because the EEG signal is a continuous time sequence. We experimented with different sizes (19, 17, 15, and 13) of convolutional kernels, and chose 17 and 15 size convolution kernels kernels by considering the computational cost and model performance. The kernels with 12 stride of all convolutional layers are intended to reduce data dimensions and retain useful information. In addition, max-pooling layers are used to reduce the dimension of the feature maps while preserving the main valid features. Also, we applied a batch normalization layer after each of the first three max-pooling layers. These layers can normalize the sensor channels in order to reduce internal covariate shift, accelerate the training process, and improve model training accuracy and generalization ability. The principle of batch normalization is as follows:

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}}$$  \hspace{1cm} (2)

$$y_i = \gamma \hat{x}_i + \beta$$ \hspace{1cm} (3)

Equation (2) is to normalize the training data of the batch, where $B$ represents a small batch that contains $m$ examples. $\mu_B$ and $\sigma_B^2$ represent the mean and variance of $B$, respectively, and $\epsilon$ is to avoid the tiny positive number used when the divisor is zero.

Equation (3) is to perform scale transformation and shift: multiply $\hat{x}_i$ by $\gamma$ to adjust the value, plus $\beta$ to increase the shift to get $y_i$, where $\gamma$ is the scale factor and $\beta$ is the translation factor. $\gamma$ and $\beta$ are learned by the network during training. They solve the problem that the normalized $\hat{x}_i$ is basically limited to a normal distribution, which reduces the expressive ability of the network.

### 3.3 LSTM

Recurrent Neural Network (RNN) is a neural network used to process sequence data. Compared with the general neural network, it can process the data of the sequence change. LSTM is a special kind of RNN, mainly to solve the problem of gradient disappearance and gradient explosion during long sequence training. Compared to ordinary RNN, LSTM can perform better in longer sequences.

We applied a LSTM layer to learn temporal information such as sleep stage scoring rules. Formally, suppose $x^t (t = 1, 2, ..., N)$ has N features from the CB, the LSTM learning process is defined as follows:

$$c^t = \sigma' \odot c^{t-1} + \varepsilon' \odot z$$ \hspace{1cm} (4)

$$h^t = \varepsilon'' \odot \tanh(c^t)$$ \hspace{1cm} (5)

$$y^t = \sigma(W'h^t)$$ \hspace{1cm} (6)
The states $z^f, z^i, z^o$ and $z$ obtained by using the current input $x^t$ of the LSTM and the $h^{t-1}$ passed down from the previous state to splice and train. Multiply the splicing vector by the weight matrix, and then utilize an activation function to convert it to a value between 0 and 1 to get $z^f, z^i$ and $z^o$ as a gated states. $z$ is to convert the splicing vector into a value between -1 and 1 through a tanh activation function as input data. $\odot$ is Hadamard Product, that is, the corresponding elements in the operation matrix are multiplied.

There are three main stages inside LSTM: (1) $z^f$ as the forgetting gate to control the $c^{t-1}$ of the previous state which needs to be left and forgotten. (2) $z^i$ as the input gate to focus on recording important parts of input $x^t$. The current input content is represented by the previously calculated $z$. The results obtained in the above two steps are added together to get the $c^t$ transmitted to the next state (see (4)). (3) $z^o$ as the output gate to determine which will be regarded as output of the current state. The output $y^t$ obtained from the change of $h^t$ is used as the input of the fully connected layers, where, $W'$ is weight matrix and $\sigma$ is sigmoid activation function.

### 3.4 Loss functions and the optimization

We experimented with four different loss functions (categorical cross-entropy, categorical hinge, logcosh and poisson) and performed nimi-batch training. Let $m$ denotes the number of examples of the mini-batch, $y_i$ represents the one-hot encoded target classes and $p_i$ is probability distribution of the sample prediction output by the soft-max layer.

The expression of the loss functions are follow:

\[
L_c(p_i, y_i) = -\frac{1}{m} \sum_{i=1}^{m} y_i \log p_i
\]

\[
L_h(p_i, y_i) = \frac{1}{m} \sum_{i=1}^{m} \max(0, 1 - y_i \cdot p_i)
\]

\[
L_l(p_i, y_i) = -\frac{1}{m} \sum_{i=1}^{m} \log(\cosh(y_i - p_i))
\]

\[
L_p(p_i, y_i) = \frac{1}{m} \sum_{i=1}^{m} (p_i - y_i \log(p_i))
\]

Categorical cross-entropy loss function $L^c$ introduces a balance parameter, which can balance positive and negative samples, and it is a convex optimization function. Categorical hinge loss function $L^h$ is used for max margin classification. Logcosh loss function $L^l$ has all the advantages of Huber loss without being affected too much by outliers. Poisson loss function $L^p$ uses an iterative method to perform weighted least squares to minimize the loss to fit the model.

Classically, the backpropagation algorithm was applied to get the gradient. For optimization, we selected Adam optimizer with parameters ($l_r = 1 \times 10^{-5}, \beta_1 = 0.9, \beta_2 = 0.999$). In addition, L2 regularization is used to avoid over-fitting.
3.5 Automatic sleep staging based on edge AI

Sleep staging on the cloud server requires a lot of network resources and a long response time, and there is a risk of disclosing user privacy. Therefore, we develop a lightweight automatic sleep staging method for children using single-channel EEG based on edge AI. The schematic diagram is shown in Figure 3. The trained model is deployed to the intelligent terminals. Therefore, the collected EEG recordings are recorded on the intelligent terminals for automatic sleep staging. Then, the results of sleep staging and hypnogram will be uploaded to the cloud server for further analysis by the physicians. Finally, the sleep disease diagnosis reports and treatment opinions are transmitted to the intelligent terminals.

Intelligent terminal devices (e.g. personal computers, smart phones and portable wearable devices etc.) have become part of people’s daily lives. After model training, it can be easily implemented on daily intelligence terminals, and only the sleep staging results need to be uploaded, which greatly saves network resources and protects user privacy to a certain extent.

4 Experiments

In this section, we firstly introduced our CS dataset and public dataset Sleep-EDFX. Then, we provide description of data processing and experiments implementation details. Finally, we present the experimental results on two datasets and analysis results.

![Figure 3 Schematic diagram of automatic sleep staging method based on edge AI](image.png)
4.1 Dataset

4.1.1 CS dataset

Our CS dataset was collected from Department of Respiratory Medicine, Beijing Children’s Hospital, Capital Medical University, National Center for Children’s Health, China. Table 2 shows demographic information of children in our CS dataset. The dataset contains 26 PSG recordings of the kids from 2 to 12 years old (8 females and 18 males). Among them, 15 subjects have OSA and 11 subjects are healthy. According to the sleep time of each subject, the collected multi-channel physiological signal is 8 to 11 hours long from the evening to the next morning. The PSG recordings of these 26 subjects contained six EEG channels: F3-M2, F4-M1, C3-M2, C4-M1, O1-M2, O2-M1. The sampling frequency of EEG signal is 256 Hz. For PSG recordings, every 30 s time interval corresponds to a label, representing one of five sleep stages (e.g., W, REM, N1, N2 and N3). The labels were provided by technicians according to the AASM sleep scoring rules. The distribution of various sleep stages in our CS dataset is shown in Figure 4(a).

| Subject ID | Sex | OSA | Night (light off) |
|------------|-----|-----|------------------|
| No.1       | Male| No  | 19:54:10 - 5:00:11 |
| No.2       | Male| No  | 21:12:03 - 6:00:08 |
| No.3       | Male| No  | 19:40:53 - 5:15:05 |
| No.4       | Female| No | 19:55:13 - 5:33:34 |
| No.5       | Female| No | 19:13:08 - 5:01:40 |
| No.6       | Female| No | 19:49:56 - 6:10:02 |
| No.7       | Male| No  | 19:06:17 - 5:39:33 |
| No.8       | Male| No  | 19:20:18 - 6:31:44 |
| No.9       | Male| Yes | 20:49:13 - 6:03:50 |
| No.10      | Female| Yes | 19:58:36 - 5:02:55 |
| No.11      | Male| Yes | 20:16:17 - 5:39:57 |
| No.12      | Male| Yes | 19:24:27 - 7:15:14 |
| No.13      | Male| Yes | 19:24:57 - 5:23:20 |
| No.14      | Male| Yes | 19:52:34 - 5:04:00 |
| No.15      | Male| No  | 19:25:53 - 5:57:14 |
| No.16      | Male| Yes | 20:09:43 - 5:02:00 |
| No.17      | Male| Yes | 20:33:56 - 5:35:26 |
| No.18      | Male| Yes | 19:52:59 - 5:22:21 |
| No.19      | Male| Yes | 20:52:59 - 5:33:19 |
| No.20      | Female| Yes | 20:39:03 - 5:09:51 |
| No.21      | Male| No  | 21:07:54 - 5:25:05 |
| No.22      | Female| Yes | 20:02:50 - 5:31:06 |
| No.23      | Male| Yes | 19:39:34 - 5:00:08 |
| No.24      | Male| No  | 20:43:08 - 5:08:19 |
| No.25      | Female| Yes | 20:34:24 - 6:32:02 |
| No.26      | Female| Yes | 21:48:58 - 5:26:42 |
4.1.2 Sleep-EDFX dataset

The sleep-EDFX [9, 19] is a well-known publicly available dataset which contains 197 whole-night PSG sleep recordings, containing EEG, EOG, chin EMG, and event markers. There are 153 PSG recordings from Sleep Cassette Study (SC), and 44 PSG recordings from Sleep Telemetry Study (ST). All PSG recordings contained two EEG channels: Fpz-Cz and Pz-Oz. The sampling frequency of EEG signal is 100 Hz. Corresponding hypnograms (sleep patterns) were manually scored by well-trained technicians according to the R&K rules. The distribution of different sleep stages in the sleep-EDFX dataset is shown in Figure 4(b). In our study, we decided to randomly select 20 PSG recordings in the ST study. We considered 5-stage sleep staging according to the AASM rules. In the R&K rules, all sleep recordings are divided into 6 stages [22]. They include W, REM and NREM, and the NREM includes Stage 1 (S1), Stage 2 (S2), Stage 3 (S3) and Stage 4 (S4). S1 and S2 are light sleep stages, and S3 and S4 are deep sleep stages. In AASM rules, the NREM includes N1 (transition stage), N2 (light sleep) and N3 (deep sleep), which respectively replace R&K stages W, REM, S1, S2, S3, and S4. The relationship between the R&K and the AASM rules is shown in Figure 5.

![Figure 4](image1.png)  Distribution of various sleep stages in the two dataset

![Figure 5](image2.png)  The relationship between the R&K and the AASM rules
4.2 Data processing

In our CS dataset, we selected the F4-M1 channel EEG signal to study automatic sleep staging taking 30 s EEG epochs as input. We also verified the effect of other channels of EEG for children’s automatic sleep staging, and the F4-M1 channel EEG got the best results. The experimental results of the six EEG channels in the CS dataset are shown in Table 3. In order to extract 30 s EEG epoch from single-channel EEG signal, we performed two steps in data processing:

- Dividing the continuous raw single-channel EEG into a sequence of 30 s epochs and assigning a label to each epoch according to the annotation file.
- Normalizing 30 s EEG epoch such that each one has a zero mean and unit variance.

The raw single-channel EEG recordings is a continuous time sequence of about 10 hours. After the first step of data processing performed, the EEG recordings of each subject can be divided into 900 to 1300 30 s epochs according to the length of sleep. So that, the raw single-channel EEG was processed into a dataset that can be used for model training and prediction. When collecting the clinical sleep EEG recordings of children, the noises generated by various reasons will affect the quality of the original signal. Normalization operation can effectively reduce the impact of these noises.

For the sleep-EDFX dataset, we perform the above data processing on Fpz-Cz and Pz-Oz channel EEG, respectively. Then, joining the 30 s Pz-Oz channel EEG with the corresponding 30 s Fpz-Cz channel EEG so that the sampling frequency of the 30 s EEG is 200 Hz as input. There are two reasons for this operation: (1) In this way, more valid information can be retained as input, and our model only needs to change the data dimension of the input layer. (2) In AASM manual, the sampling frequency of the EEG signal should be between 200Hz and 500Hz, where 500Hz is the ideal value and 200Hz is the minimum value.

4.3 Experiments implementation

For the two data sets, we randomly select a small number of EEG recordings as the test set, and the others as the training set. The specific division ratio is shown in Sections 4.4 and 4.5. We used 10-fold cross-validation to train our model with 100 epochs, and batch size is 256. The performance of the model is evaluated according to the following test criteria: confusion matrix, accuracy and F1 score. The accuracy can intuitively reflect the proportion of correct sleep stages in the prediction results. The confusion matrix can reflect the

| Channel | Accuracy | F1-score |
|---------|----------|----------|
| F3-M2   | 80.42    | 73.59    |
| F4-M1   | 83.06    | 76.50    |
| C3-M2   | 79.74    | 73.87    |
| C4-M1   | 81.40    | 74.41    |
| O1-M2   | 72.93    | 67.85    |
| O2-M1   | 73.68    | 65.34    |

Table 3 The experimental results of the six EEG channels in the CS dataset
prediction results of all sleep stages, where the diagonal elements are the accuracy of each sleep stage. The F1-score calculated from the confusion matrix can comprehensively represent the performance of the model.

Our experimental models were implemented using Keras in the Tensorflow framework under the Python environment. Our experiments were conducted by a desktop PC equipped with Intel Intel i7-8700K CPU, 64 GB RAM and a NVIDIA GeForce GTX 1080Ti GPU.

4.4 Performance evaluation on the our CS dataset

For 26 subjects in our CS dataset, we randomly selected 5 subjects as the test set, and the others as the training set. The training and validation accuracy with different loss functions is shown in Figure 6. For multi-class problems, the most commonly used loss function is categorical cross-entropy, due to the loss is only proportional to the difference between the output value and the true value, the convergence is fast, and it can be optimized through the backpropagation algorithm. In subfigure (a), training accuracy keeps increasing with the increase of epochs, and it reaches about 94% after 100 epochs, but the validation accuracy does not continue to increase when it reaches about 86%. Theoretically, since the optimization of the hinge loss is less than a certain gap distance, the optimization will stop, but the cross-entropy loss is always optimized. Therefore, under normal circumstances, the effect of the cross-entropy loss is better than that of

![Figure 6](image_url)
the hinge loss. The training accuracy is only about 87% after 100 epochs in subfigure (b). However, the validation accuracy can reach up to about 89%, and it is consistently higher than the validation accuracy in subfigure (a) after 20 epochs. However, it can be seen from Figure 8 that all evaluation metrics of categorical hinge loss function are lower than others. Subfigure (c) is similar to subfigure (b), but the training accuracy is a little higher after 100 epochs, and the validation accuracy curve seems to change more frequently. Subfigure (d) is similar to subfigure (a), but the training accuracy is a little lower after 100 epochs.

After training the model, we used the test set to test the performance of the model with different loss functions. The confusion matrix with sleep staging results for different sleep stages is shown in the Figure 7. In subfigure (a), the sleep staging performance of the stage N2 and N3 is satisfactory, and the stage W and REM is acceptable, but stage N1 got a bad result. The accuracy of stage W, N3 and REM in subfigure (b) is better than others, but the stage N1 and N2 is the worst. In subfigure (c), all of the stage N2, N3 and REM get an accuracy of about 90%. The accuracy of stage N3 and REM is more than 90%, but the stage W and N2 is only a little over 80%.

Figure 7 The confusion matrix with different loss functions
It is easy to see in Figure 7 that in all models trained with different loss functions, the results of stage N1 are very poor. A large number of stage N1 are mistakenly classified as stage W, N2 and REM. Stage N1 is a transitional stage in sleep, and the EEG features of stage N1 are not obvious. It is also difficult for well-trained technicians to classify stage N1 accurately [4, 28]. Similarly, for the REM stage, staging errors are mainly mistakenly regarded as N1 stage. Some the of stage W epochs are mistakenly classified as the stage N1 and N2, a small amount of stage N2 epochs considered stage N1 and N3, and a few stage N3 epochs are regarded as stage N2. There is an interesting phenomenon here: most of the incorrectly divided sleep stages correspond to adjacent stages of the correct sleep stage. The sleep stages are contiguous in the sleep cycle, therefore each sleep stage may contain patterns similar to the adjacent stage. In addition, this phenomenon may also be caused by mislabeling of adjacent sleep stages by technicians. For all loss functions, the accuracy of stage N3 and REM is satisfactory, and stage W is acceptable. Stage N2 accounts for a large proportion of the sleep recordings (see Figure 4(a)), therefore the accuracy of the stage N2 has a greater impact on the overall accuracy. Therefore, How to improve the sleep staging accuracy of the stage N1 and N2 is the focus of our further work.

The overall accuracy, precision, recall and F1-score for different loss functions is shown in Figure 8. On the whole, the performance of models trained with different loss functions is not much different, an accuracy of 82.15%, 80.13%, 83.06% and 81.96% is for categorical cross-entropy, categorical hinge, logcosh and poisson, respectively. Just observe Figure 7 (b), it seems that the performance of sleep staging is very good. in fact, the evaluation metrics of categorical hinge is the lowest due to cannot recognize stage N2 well. However, the accuracy of its stage W exceeds other loss functions by about 6.5%. The logcosh has the highest accuracy due to outliers of raw EEG have little effect on it, and poisson has the best precision, recall and F1-score. For categorical cross-entropy, The accuracy is close to logcosh, and other evaluation metrics are close to poisson. Different loss functions have their own advantages and disadvantages in sleep staging. In order to further improve the performance of automatic sleep staging for children, we consider proposing a new loss function based on the above four loss functions. An easy idea is to perform a weighted average of the four loss functions. The new loss function is follow: 

\[
\text{Loss} = w \cdot [L_c, L_h, L_l, L_p]^T
\]

where \(w\) represents a weight vector. Unfortunately,
the results obtained are not ideal. We will further study the principles of these loss functions and hope to propose a loss function that is more suitable for children’s automatic sleep staging.

4.5 Comparative experiment on the sleep-EDFX dataset

For 20 PSG recordings in the sleep-EDFX dataset we randomly selected 4 recordings as the test set (20%), and the others as the training set (80%). According to AASM rules, the performance of the proposed method on the sleep-EDFX dataset is shown in Table 4, and the performance comparison between our CSleepNet and some state-of-the-art methods on the Sleep-EDFX dataset is shown in Table 5. There is high variance in the experimental implementation details (e.g., the number of EEG recordings, methods, used EEG channel) observed in different literature, but most past works used accuracy as their main evaluation metric. Therefore, we select the accuracy as main evaluation metric for this comparative experiment for five stages sleep staging by the Fpz-Cz and Pz-Oz channel EEG signals of the Sleep-EDFX dataset.

In the Table 5, the first four articles are based on traditional machine learning methods, the others are based on deep learning. It can be seen that the sleep staging accuracy with hand-crafted features are higher than 91%. Although the accuracy of deep learning methods
is not as good as machine learning, the results are also satisfactory. Without adjusting our model for the sleep-EDFX dataset, we achieved an accuracy of 86.41%. Although the accuracy rate is a little lower compared with the top methods, it is enough to show that our method is effective. The gap in performance of CSleepNet in these two data sets, we consider, was caused by the following reasons. (1) In the process of sleep monitoring, children are more sensitive to the monitoring equipment, and children will feel discomfort due to the monitoring equipment, so noises such as artifacts appear and the equipment may even fall. (2) For our CS dataset, most subjects have varying degrees of OSA. For Sleep-EDFX dataset, most subjects are healthy, and others just suffer from mild sleep disorders. OSA leads EEG changes during sleep more complicated. (3) We used two channel EEG on Sleep-EDFX dataset. Using multi-channel data may be more beneficial for sleep staging, but it requires more computing resources, and the use of single-channel EEG for sleep staging can also achieve satisfactory results. Therefore, in order to deploy the model more conveniently on home computers, mobile phones, smart bracelets and other smart devices, single-channel EEG is still the best choice at present. The experimental results show that our method has great potential. The realization of children’s automatic sleep staging based on edge AI is of great significance to the health of children.

5 Conclusions

Based on edge AI, we combined 1D-CNN and LSTM to propose lightweight automatic sleep staging for children using single-channel EEG. The experimental results show that a single-channel EEG and the CSleepNet can be used to sleep staging without any feature extraction stage, and its performance is satisfactory. This has an advantage that the model can be trained to learn the features that are most suited to the sleep staging for children. Different loss functions have their own advantages in different stages of sleep staging. Training the CSleepNet model requires a lot of time and hardware equipment with sufficient performance, but once the model training is completed, the prediction is relatively cheap, and can be carried out on personal computers, mobile phone and portable wearable devices.

In future work, we will further improve the automatic sleep staging method for children using single-channel EEG based on edge AI. The sleep staging algorithm on the edge device side has been implemented, therefore we will concentrate on designing the analysis algorithm based on the sleep staging results and the hypnogram, and deploy it to the cloud server. The algorithm is expected to provide sleep quality reports, identify respiratory events, and assist the diagnosis of sleep diseases. In addition, we will also collect more sleep recordings of children with reliable annotations, and conduct more experiments based on a larger amount of data and propose a loss function that is more suitable to further improve children’s sleep staging model.

Acknowledgements This work was supported in part by the National Natural Science Foundation of China under Grant 62172340, in part by the Natural Science Foundation of Chongqing under Grant cstc2021jcyj-msxmX0041, in part by the Young and Middle-aged Senior Medical Talents Studio of Chongqing under grant ZQNYXGDRCGZS2021002, and in part by the Introduced Talent Program of Southwest University under Grant SWU020008. The corresponding author is Yuan Zhang.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.
References

1. Agarwal, R., Gotman, J.: Computer-assisted sleep staging. IEEE transactions on bio-medical engineering 48(12), 1412–1423 (2001). https://doi.org/10.1109/10.966600
2. Agarwal, A., Sharma, S., Kumar, V., Kaur, M.: Effect of E-Learning on Public Health and Environment During COVID-19 Lockdown. Big Data Mining and Analytics 4(2), 104–115 (2021)
3. Alickovic, E., Subasi, A.: Ensemble SVM Method for Automatic Sleep Stage Classification. IEEE Transactions on Instrumentation and Measurement 67(6), 1258–1265 (2018). https://doi.org/10.1109/TIM.2018.2799059
4. Berry, R.B., Quan, S.F., Abreu, A.R., et al.: The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications (Version 2.6), american academy of sleep medicine (2020)
5. Bevans, K.B., Meltzer, L.J., Anna, D.L.M., Kratchman, A., Vil, D., Forrest, C.B.: Qualitative development and content validation of the PROMIS pediatric sleep health items. Behav. Sleep Med. 657–671 (2018). https://doi.org/10.1080/15402002.2018.1461102
6. Boostani, R., Karimzadeh, F., Nami, M.: A comparative review on sleep stage classification methods in patients and healthy individuals. Computer Methods and Programs in Biomedicine 140, 77–91 (2017). https://doi.org/10.1016/j.cmpb.2016.12.004
7. Estrada, E., Nazeran, H., Nava, P., Behbehani, K., Burk, J., Lucas, E.: EEG feature extraction for classification of sleep stages. In: The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 196–199 (2004). https://doi.org/10.1109/EMBS.2004.1403125
8. Fraiwan, L., Alkhodari, M.: Investigating the use of uni-directional and bi-directional long short-term memory models for automatic sleep stage scoring. Inform. Med. Unlocked 20, (2020). https://doi.org/10.1016/j.imu.2020.100370
9. Goldberger, A., Amaral, L., Glass, I., et al.: PhysioBank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. Circulation Online 101(23), 215–220 (2000)
10. Grigg-Damberger, M., Gozal, D., Quan, S.F., Rosen, C.L.: The visual scoring of sleep and arousal in infants and children. Journal of Clinical Sleep Medicine Jcsm Official Publication of the American Academy of Sleep Medicine 3(2), 201–240 (2007)
11. Hanan, E.S.: Sleep physiology and sleep disorders in childhood. Nature and Science of Sleep 3, 101–114 (2011). https://doi.org/10.2147/NSS.S22839
12. Hassan, A.R., Bhuiyan, M.: A decision support system for automatic sleep staging from EEG signals using tunable Q-factor wavelet transform and spectral features. Journal of Neuroscience Methods 271, 107–118 (2016). https://doi.org/10.1016/j.jneumeth.2016.07.012
13. Hassan, A.R., Bhuiyan, M.: An automated method for sleep staging from EEG signals using normal inverse Gaussian parameters and adaptive boosting. Neurocomputing 219, 76–87 (2016). https://doi.org/10.1016/j.neucom.2016.09.011
14. Hassan, A.R., Bhuiyan, M.: A decision support system for automated identification of sleep stages from single-channel EEG signals. Knowledge-Based Systems 128, 115–124 (2017). https://doi.org/10.1016/j.knosys.2017.05.005
15. He, Q., Cui, G., Zhang, X., Chen, F., Deng, S., Jin, H., Yang, Y.: A Game-Theoretical Approach for User Allocation in Edge Computing Environment. IEEE Transactions on Parallel and Distributed Systems 31(3), 515–529 (2020). https://doi.org/10.1109/TPDS.2019.2938944
16. Hus, Y., Yang, Y., Wang, J., Hsu, C.: Automatic sleep stage recurrent neural classifier using energy features of EEG signals. Neurocomputing 104(1), 105–114 (2013). https://doi.org/10.1016/j.neucom.2012.11.003
17. Huang, Z., Lin, K.J., Tsai, B.L., Yan, S., Shih, C.: Building edge intelligence for online activity recognition in service-oriented IoT systems. Future Generation Computer Systems 87, 557–567 (2018). https://doi.org/10.1016/j.future.2018.03.003
18. Huang, H., Yang, L., Wang, Y., Xu, X., Lu, Y.: Digital Twin-driven online anomaly detection for an automation system based on edge intelligence. Journal of Manufacturing Systems 59, 138–150 (2021). https://doi.org/10.1016/j.jmsy.2021.02.010
19. Kemp, B., Zwiderman, A.H., Tuk, B., Kamphuisen, H.A.C., Obery, J.J.L.: Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the EEG. IEEE-BME 47(9), 1185–1194 (2000). https://doi.org/10.1109/667928
20. Khandewale, A., Gohokar, V., Nawandar, P.: Edge intelligence-based object detection system using neural compute stick for visually impaired people. In: Senjyu, T., Mahalle, P.N., Perumal, T., Joshi, A. (eds) Information and Communication Technology for Intelligent Systems. ICTIS 2020. Smart Innovation, Systems and Technologies, 195, 433–439. Springer, Singapore (2021). https://doi.org/10.1007/978-981-15-7078-0_41
21. Liu, L., Lu, Z., Wang, L., Chen, X., Wen, X.: Large-volume data dissemination for cellular-assisted automated driving with edge intelligence. J. Netw. Comput. Appl. 155(2) (2020). https://doi.org/10.1016/j.jnca.2020.102535

22. Reschttchaffen, A., Kales, A.: A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects. Electroencephalography and Clinical Neurophysiology 26(2), 644 (1969). https://doi.org/10.1016/0140-6736(69)90021-2

23. Rincon, J.A., Julian, V., Carrascosa, C.: Towards the edge intelligence: robot assistant for the detection and classification of human Emotions. In: De La Prieta, F., et al. (eds) Highlights in Practical Applications of Agents, Multi-Agent Systems, and Trust-worthiness. The PAAMS Collection. PAAMS 2020. Communications in Computer and Information Science, 1233, 31-41, Springer, Cham (2020). https://doi.org/10.1007/978-3-030-51999-5_3

24. See A.R., Liang C.K.: A study on sleep EEG using sample entropy and power spectrum analysis. In: 2011 Defense Science Research Conference and Expo (DSR), pp. 1–4 (2011). https://doi.org/10.1109/DSR.2011.6026802

25. Shi, W., Cao, J., Zhang, Q., Li, Y., Xu, L.: Edge Computing: Vision and Challenges. IEEE Internet of Things Journal 3(5), 637–646 (2016). https://doi.org/10.1109/JIOT.2016.2579198

26. Song, Z., Cao, Z., Li, Z., Wang, J., Liu, Y.: Inertial Motion Tracking on Mobile and Wearable Devices: Recent Advancements and Challenges. Tsinghua Science and Technology 26(5), 692–705 (2021)

27. Sors, A., Bonnet, S., Mirek, S., Vercueil, L., Payen, J.: A convolutional neural network for sleep stage scoring from raw single-channel EEG. Biomedical Signal Processing and Control 42, 107–114 (2018). https://doi.org/10.1016/j.bspc.2017.12.001

28. Stepnowsky, C., Levendowski, D., Popovic, D., Ayappa, I., Rapoport, D.M.: Scoring accuracy of automated sleep staging from a bipolar electrooculographic recording compared to manual scoring by multiple raters. Sleep Medicine 14(11), 1199–1207 (2013). https://doi.org/10.1016/j.sleep.2013.04.022

29. Supratak, A., Hao, D., Chao, W., Guo, Y.: DeepSleepNet: A Model for Automatic Sleep Stage Scoring Based on Raw Single-Channel EEG. IEEE Transactions on Neural Systems and Rehabilitation Engineering 25(11), 1998–2008 (2017). https://doi.org/10.1109/TNSRE.2017.2721116

30. Uddin, M.Z., Nilsson, E.G.: Emotion recognition using speech and neural structured learning to facilitate edge intelligence - ScienceDirect. Eng. Appl. Artif. Intell. 94 (2020). https://doi.org/10.1016/j.engappai.2020.103775

31. Zhang, J., Wu, Y.: Competition convolutional neural network for sleep stage classification. Biomed. Signal Process. Control 64 (2021). https://doi.org/10.1016/j.bspc.2020.102318

32. Zhang, J., Yan, W.: A New Method for Automatic Sleep Stage Classification. IEEE Transactions on Biomedical Circuits and Systems 11(5), 1097–1110 (2017). https://doi.org/10.1109/TBCAS.2017.2719631

33. Zhang, Z., Cong, X., Feng, W., Zhang, H., Fu, G., Chen, J.: WAEAS: An Optimization Scheme of EAS Scheduler for Wearable Applications. Tsinghua Science and Technology 26(1), 72–84 (2021)

34. Zhang, C., Liu, S., Han, F., Nie, Z., Lo, B., Zhang, Y.: Hybrid manifold-deep convolutional neural network for sleep staging. Methods (2021). https://doi.org/10.1016/j.ymeth.2021.02.014

35. Zhen, P., Han, Y., Dong, A., Yu, J.: CareEdge: A Lightweight Edge Intelligence Framework for ECG-Based Heartbeat Detection. Procedia Computer Science 187, 329–334 (2021). https://doi.org/10.1016/j.procs.2021.04.070

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
Authors and Affiliations

Liqiang Zhu1 · Changming Wang2,3 · Zhihui He4 · Yuan Zhang1

Changming Wang
superwcm@163.com

Zhihui He
hezhihui726@sina.com

1 College of Electronic and Information Engineering, Southwest University, Chongqing 400715, China
2 Department of Neurosurgery, Xuanwu Hospital, Capital Medical University, Beijing 100053, China
3 Brain-inspired Intelligence and Clinical Translational Research Center, Beijing 100176, China
4 Department of Pediatric Respiration, Chongqing Ninth People’s Hospital, Chongqing 400700, China