Automatic Identification of Insomnia Based on Single-Channel EEG Labelled With Sleep Stage Annotations

BUFANG YANG AND HONGXING LIU
School of Electronic Science and Engineering, Nanjing University, Nanjing 210008, China
Corresponding author: Hongxing Liu (njhxliu@nju.edu.cn)

This work was supported by the Fundamental Research Funds for the Central Universities under Grant 021014380163.

ABSTRACT Monitoring single-channel EEG is a promising home-based approach for insomnia identification. Currently, many automatic sleep stage scoring approaches based on single-channel EEG have been developed, whereas few studies research on automatic insomnia identification based on single-channel EEG labelled with sleep stage annotations. In this paper, we propose a one-dimensional convolutional neural network (1D-CNN) model for automatic insomnia identification based on single-channel EEG labelled with sleep stage annotations, and further investigate the identification performance based on different sleep stages EEG epochs. Single-channel EEG on 9 insomnia patients and 9 healthy subjects was used in this study. We constructed 4 subdatasets from EEG epochs based on the sleep stage annotations: All sleep stage dataset (ALL-DS), REM sleep stage dataset (REM-DS), light sleep stage dataset (LSS-DS), and SWS sleep stage dataset (SWS-DS). Subsequently, 4 subdatasets were fed into our 1D-CNN. We conducted experiments under intra-patient and inter-patient paradigms, respectively. Our experiments demonstrated that our 1D-CNN leveraging 3 subdatasets composed of REM, LSS and SWS epochs, respectively, achieved higher average accuracies in comparison with baseline methods under both intra-patient and inter-patient paradigms. The experimental results also indicated that amongst all the sleep stages, 1D-CNN leveraging REM and SWS epochs exhibited the best insomnia identification average accuracies in intra-patient paradigm, which are 98.98% and 99.16% respectively, whereas no statistically significant difference was found in inter-patient paradigm. For automatic insomnia identification based on single-channel EEG labelled with sleep stage annotations, 1D-CNN model introduced in this paper could achieve superior performance than traditional methods.

INDEX TERMS Convolutional neural networks, insomnia, inter-patient paradigm, intra-patient paradigm, single-channel electroencephalogram (EEG), sleep stage, sleep data analysis.

I. INTRODUCTION
Sleep is a fundamental physiological activity, which plays a crucial role in physical and mental health for human body [1]. Insomnia is a sleep disorder that is prevalent in adults [2]. In the clinical practice, clinicians diagnose insomnia through sleep questionnaires, polysomnography (PSG) monitoring of the patients, and the diagnostic criteria for insomnia released by American Academy of Sleep Medicine (AASM) [3], [4]. However, the subjectivity of the sleep questionnaires and the first-night effect of the PSG recordings make the insomnia diagnose a time-consuming, expensive and subjective process, which is unsuitable for home-usage [5], [6].

Home-based sleep monitoring, which is a hot research area, has many approaches: (1) smart mats based on piezo-electric and pressure sensors, (2) electrocardiogram (ECG) and pulse wave, (3) electroencephalogram (EEG) [7]. Various algorithms have been proposed to tackle the problem of automatic sleep disorder detection based on the above approaches. Hassan et al. [8] extracted statistical features in the tunable-Q factor wavelet domain and classified obstructive sleep apnea (OSA) by random under sampling boosting (RUSBoost) classifier based on single-lead ECG. Heyat et al. [9] leveraged the power spectral density (PSD) features and decision tree classifier for sleep bruxism.
Current automatic insomnia identification methods based on single-channel EEG require experience-based handcrafted features to train a traditional classifier. Aydin et al. [12] extracted the 10-dimensional singular spectrum features of the sleep EEG signal, and then fed the features into a single hidden layer artificial neural network (ANN) for insomnia identification. Hamida et al. [13] extracted spectral and Hjorth’s parameters features, and applied principal component analysis (PCA) for dimension reduction. The first principal component was then classified according to the set threshold. Shahin et al. [14] extracted statistical, temporal and spectral features of EEG signals, and leveraged deep neural network (DNN) for automatic insomnia identification. Zhang et al. [15] proposed an insomnia identification method based on temporal, spectral, nonlinear features and random forest (RF) classifier. Therefore, till date, most of the existing work for automatic insomnia identification task was based on hand-crafted feature extraction and traditional machine learning algorithms.

Convolutional neural networks (CNN) do not need to define features manually, which can overcome the limitation of the handcrafted features that are limited by prior knowledge. This advantage makes CNN gain much attention in biomedical engineering field [16], [17]. Since most of the existing work for automatic insomnia identification task was based on hand-crafted features and traditional machine learning algorithms, an end-to-end one-dimensional CNN (1D-CNN) model based on single-channel EEG signal is investigated in this study.

Sleep is a cyclical process, which is composed of three main stages: rapid eye movement, light sleep and deep sleep [18]. According to the Rechtschaffen and Kales (R&K) rules [19], an overnight EEG sleep signal is divided into 30-second epochs, and each epoch is categorized as wake, rapid eye movement (REM) and non-rapid eye movement (NREM) stage, which is further divided into S1, S2, S3 and S4 stages. Table 1 summarizes the sleep stage scoring criteria and description. Many approaches for automatic sleep stage scoring based on single-channel EEG have been developed, which can achieve high identification performance. Hassan and Bhuiyan [20] extracted various statistical features in ensemble empirical mode decomposition (EEMD) and leveraged RUSBoost classifier for sleep stage classification. Jiang et al. [21] leveraged 151-dimensional time and frequency domain features, RF classifier and proposed hidden Markov model (HMM) based rule refinement to identify sleep stages. Supratak et al. [22] proposed a deep learning model DeepSleepNet for automatic sleep stage classification, which contains CNN part and bidirectional-long short-term memory (LSTM) part. Chen et al. [23] proposed a deep learning model SleepStageNet including multi-scale CNN, recurrent neural networks (RNN) and conditional random field (CRF) to identify sleep stages. Therefore, in this study, we investigate the automatic insomnia identification method leveraging EEG labelled with sleep stage annotations.

Several studies have investigated the insomnia identification performance based on different sleep stages of EEG [13], [14]. Hamida et al. [13] evaluated the performance of wake, S1, S2, SWS and REM stages, finding SWS epochs have the best insomnia identification performance. Shahin et al. [14] compared the performance of all stages, NREM, S2+S3, NREM+REM stages, finding NREM+REM epochs have the best performance, whereas they did not evaluate the performance only leveraging REM epochs. At present, it is still unclear which EEG sleep stages has the best insomnia identification performance [24].

In this paper, we propose a 1D-CNN model for automatic insomnia identification leveraging single-channel EEG labelled with sleep stage annotations, and further investigate the identification performance based on different sleep stages. This is the first implementation of CNN in automatic insomnia identification task to the best of the author’s knowledge. The rest of this paper is organized as follows. Section II describes the dataset and baseline method. Section III presents our method including the general idea, data preprocessing and 1D-CNN model. Section IV presents the experiments and results. In Section V, we discuss the experimental results. Finally, Section VI concludes the paper.

### Table 1. R&K Sleep stage scoring criteria and description.

| Sleep stage | R&K classification criteria | Description |
|-------------|----------------------------|-------------|
| Wake        | Alpha power for $\geq$50% of an EEG epoch | Wakefulness |
| S1          | Vertex waves; alpha power for $<50\%$ of an EEG epoch | Light sleep; quick transition |
| S2          | Appearance of K-complexes or sleep spindles | Light sleep; low heart rate |
| S3          | Slow-wave power for 20 to 50% of an EEG epoch | Deep sleep; body self-repair |
| S4          | Slow-wave power for $\geq$50% of an EEG epoch | Deep sleep; body self-repair |
| REM         | Episodic rapid eye movements; reduced submental EMG activity; mixed-frequency EEG activity | Reduced movements; dreams; rapid eye movements |
TABLE 2. Subject information in this study.

| Subject name | Start time | End time | No. of epochs (C4-A1 channel) | fs (Hz) |
|--------------|------------|----------|-------------------------------|--------|
| Healthy-1    | 22:09:33   | 07:42:33 | 30 33 513 136 186 239 512     |        |
| Healthy-2    | 22:19:06   | 06:38:36 | 143 141 368 83 114 151 512    |        |
| Healthy-3    | 23:06:12   | 07:26:12 | 136 49 348 112 168 188 512    |        |
| Healthy-4    | 23:36:37   | 07:01:37 | 223 18 417 64 80 209 100      |        |
| Healthy-5    | 22:49:48   | 07:13:18 | 10 49 414 134 169 232 512     |        |
| Healthy-10   | 23:24:52   | 06:34:22 | 65 2 63 17 187 17 512         |        |
| Healthy-11   | 22:37:16   | 07:23:16 | 56 6 58 36 184 22 512         |        |
| Healthy-12   | 15:14:22   | 23:28:52 | 31 5 424 74 138 298 100      |        |
| Healthy-14   | 22:15:32   | 06:19:02 | 12 7 322 146 131 149 200     |        |
| Insomnia-1   | 22:30:28   | 07:07:58 | 121 34 496 131 0 134 256      |        |
| Insomnia-2   | 18:23:38   | 08:22:38 | 837 5 455 145 0 233 512       |        |
| Insomnia-3   | 22:00:42   | 05:13:52 | 229 136 193 67 107 135 256    |        |
| Insomnia-4   | 21:34:04   | 03:40:04 | 63 6 352 61 97 154 256        |        |
| Insomnia-5   | 17:58:48   | 08:18:18 | 918 5 430 96 173 512         |        |
| Insomnia-6   | 22:37:17   | 07:25:17 | 494 62 227 64 122 88 512      |        |
| Insomnia-7   | 15:58:14   | 08:19:14 | 610 19 509 64 61 220 512      |        |
| Insomnia-8   | 22:43:04   | 05:42:34 | 232 73 268 189 0 78 512       |        |
| Insomnia-9   | 22:28:44   | 07:13:14 | 654 53 215 51 37 40 512       |        |

B. BASELINE

According to the method in [12], the trajectory matrix $X$ of EEG epoch was computed based on the phase-space reconstruction and got the covariance matrix $C = \frac{1}{N}X^TX$. Subsequently, the first 10 singular values (in descending order) of $C$ were computed by singular value decomposition (SVD). Finally, the 10-dimensional singular spectrum features were fed into ANN for insomnia identification. The ANN consisted of one hidden layer, which had 10 neurons. All layers used the sigmoid activation function. In subsequent experiments, we selected the epochs of wake, REM, S1 and S2 sleep stages in C4-A1 channel EEG for implementing this method.

Previous studies demonstrated that relative power and Hjorth parameters of EEG signal are important features for insomnia identification [29]–[31]. According to the method in [13], EEG was filtered at 6 frequency bands: delta (0.5-4Hz), theta (4-8Hz), alpha (8-12Hz), sigma (12-16Hz), beta (16-30Hz) and gamma (30-40Hz). A total of 22 features including power and Hjorth parameters were extracted. Subsequently, the dimensionality of features were decreased by PCA algorithm. Power features contain the relative power in each frequency band and their ratios. Hjorth parameters features include activity, mobility and complexity, which are defined as follows:

$$\text{activity} = \frac{\text{var}(x)}{1}$$

$$\text{mobility} = \sqrt{\frac{\text{var}(x')}{\text{var}(x)}}$$

$$\text{complexity} = \frac{\text{mobility}(x')}{\text{mobility}(x)}$$

where $x$ is the EEG epoch. According to the method in literature [13], we leveraged the epochs of SWS sleep stages in C4-A1 channel EEG for implementing this method. Then we applied PCA for dimensionality reduction and leveraged the first principal component as the final one-dimensional feature. We searched the optimal differentiate threshold for the first principal component on training dataset. Then, the optimal differentiate threshold was used in test dataset for insomnia identification. However, we considered only leveraging the first principal component feature may cause loss of information. Since RF is an ensemble learning algorithm by constructing a group of decision trees [32], which has properties of adaptability and robustness, we also used RF classifier for the above 22-dimensional features. In this way, we could further evaluate the identification performance of the 22-dimensional features. In our experiment, we found that with the number of trees increasing, the result had the highest insomnia identification accuracy when the number of trees reached 1000. When the number of trees continued to increase, the identification performance did not increase significantly, while the computation speed decreased. Therefore, the number of trees was set to 1000 in our experiment.
III. METHOD

A. GENERAL IDEA

Fig. 1 depicts the general idea of our study. Traditionally, a complete automatic insomnia identification model includes three basic sections: data preprocessing, feature extractor and classifier. Since 1D-CNN integrates the feature extractor and classifier into one single algorithm, the model in this study consists of two sections: data preprocessing and 1D-CNN model. In order to investigate the superiority of 1D-CNN model, it is necessary to make a comparison with the baseline methods in [12], [13] based on the same dataset and preprocessing method. In order to investigate the generalizability of our method, we conducted experiments under both intra-patient and inter-patient paradigms.

In order to further investigate which sleep stage has better insomnia identification performance, we constructed 4 subdatasets from EEG epochs according to their sleep stage annotations. According to the AASM standard [33], we merged the S3 and S4 into single slow wave sleep (SWS) stage. The S1 and S2 were merged into light sleep stage (LSS) [18]. The 4 subdatasets are ALL-DS, REM-DS, LSS-DS, SWS-DS, which are constructed from the all sleep stage 30s epoch, i.e. the $1 \times 3840$ one-dimensional time series, 1D convolutional kernel is used in our model. The 1D convolution operation process is defined as:

$$y_i^l = f\left(\sum_{n=1}^{d} w_n^l \cdot y_{i+n}^{l-1} + b^l\right), \quad i = 1, 2, \ldots, N-d+1$$

where $y_i^l$ is the $i$th pixel of the output feature on the $l$ th layer. $w_n^l$ and $b^l$ denote the weight vector and the bias parameter of the convolutional kernel on the $l$ th layer, respectively. $d$ denotes the size of the convolutional kernel. $N$ denotes the length of input feature vector $y_{i-1}^{l-1}$. $f(\cdot)$ denotes the activation function of convolution layer.

The first and second convolution layers use large kernels with size of $1 \times 11$ and $1 \times 5$, respectively, whereas the third to fifth convolution layers use small kernels with size of $1 \times 3$. After the first, second and the fifth convolution layers, we utilized the maxpooling layer with size of $1 \times 3$ to reduce the dimension of feature maps. Subsequently, the generated feature maps of the last maxpooling layer are flattened into a one-dimensional vector. This vector are fed into the fully connected layer for binary classification, and the final identification result is obtained. We chose ReLU (Rectified Linear Unit) as the activation function, which is defined as follows:

$$f(x) = \begin{cases} x, & \text{if } x > 0 \\ 0, & \text{otherwise} \end{cases}$$

A batch normalization layer follows the first convolution layer, which can normalize the feature activations, thus reducing the internal covariate shift. The batch normalization is defined as follows [35]:

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}}$$

$$y_i = \gamma \hat{x}_i + \beta$$

where $B$ represents the mini-batch including $m$ samples, $\mu_B$ and $\sigma_B^2$ represent the mean and variance of $B$, respectively. $\epsilon$ is a constant for numerical stability. $\gamma$ and $\beta$ are the scale and shift parameters computed in the training process, respectively, which can be seen in [35] for details. The parameters associated with our CNN are given in Table 3.

B. DATA PREPROCESSING

The whole night single-channel EEG (C4-A1) was band filtered and resampled. The sleep stage annotations of epochs after the data preprocessing were unchanged.

1) BAND PASS FILTERING

Since EEG is low-frequency signal, whose frequency components are mainly concentrated in 0.5-50Hz frequency band, we designed an 80th-order band-pass FIR filter (0.5-50 Hz) based on Hamming window. The raw C4-A1 channel EEG was preprocessed by the filter for eliminating high frequency noise and direct current component.

2) RESAMPLING

As illustrated in Table 2, C4-A1 channel EEG recordings in CAP Sleep Database have different sampling frequency. Hence, all EEG recordings were resampled at 128 Hz.

C. PROPOSED MODEL

1) 1D-CNN STRUCTURE

The framework of our 1D-CNN is inspired by AlexNet [34]. We replaced the 2D convolutional kernel with 1D convolutional kernel and added batch normalization layer to our 1D-CNN, while the size of convolutional kernel remains unchanged. Fig. 2 depicts the schematic diagram of our 1D-CNN model, which consists of 5 convolution layers, 3 pooling layers and 3 fully connected layers.

Since the input of our 1D-CNN is the specified sleep stage 30s epoch, i.e. the $1 \times 3840$ one-dimensional time series, 1D convolutional kernel is used in our model. The 1D convolution operation process is defined as:
sleep stages on insomnia identification, we constructed sub-
datasets according to the sleep stage of epochs. The process
of subdatasets construction consists of two parts, selecting
specified sleep stages and overlapping.

Firstly, we selected the sleep stage of the subdataset that
needs to be constructed. Subsequently, if two consecutive
epochs of the subject had the same sleep stage, we used a
sliding window for overlapping. The overlapping time was set
to 25s. Note that the sleep stage label of the new epoch after
overlapping was unchanged. Conversely, if two consecutive
epochs had different sleep stages, we did not conduct over-
lapping. Fig. 3 shows the schematic diagram of overlapping
method in this study. With the above method, we constructed

4 subdatasets: ALL-DS, REM-DS, LSS-DS and SWS-DS. Additionally, we constructed a BSL-DS (baseline dataset)
for implementing baseline method in [12]. Fig. 4 shows the
schematic diagram of constructing subdataset in this study.

- All sleep stage dataset (ALL-DS): including all sleep
  stage epochs of the EEG recording.
- REM sleep stage dataset (REM-DS): only including
  REM epochs of the EEG recording.
- Light sleep stage dataset (LSS-DS): including S1 and
  S2 epochs of the EEG recording.
- SWS sleep stage dataset (SWS-DS): including S3 and
  S4 epochs of the EEG recording.
- Baseline dataset (BSL-DS): including wake, REM,
  S1 and S2 epochs of the EEG recording.

It should be noted that since the epochs number for each
sleep stage from each subject is different, the sizes of the
5 subdatasets constructed directly by the method mentioned
above are different. However, the size of the dataset can
greatly affect the performance of machine learning algorithm.
Therefore, we employed the following strategies to adjust the
number of subdatasets: (1) we regarded the minimum number
among the 5 types of epochs of the target subject used to
construct subdataset as the threshold. Each type of the epochs
only kept the threshold number and the rest of the epochs were
discarded. (2) if the minimum number was greater than 800,
we set the threshold as 800. We implemented those two
strategies for each subject and finally obtained the 5 sub-
datasets that had the same size as well as the same number of
epochs from each subject. For further details see Discussion.
Table 4 shows the number of epochs for the 5 subdatasets.

| Subject type | ALL-DS | REM-DS | LSS-DS | SWS-DS | BSL-DS |
|--------------|--------|--------|--------|--------|--------|
| Healthy      | 5782   | 5782   | 5782   | 5782   | 5782   |
| Insomnia     | 5970   | 5970   | 5970   | 5970   | 5970   |
| Total        | 11752  | 11752  | 11752  | 11752  | 11752  |

TABLE 3. Parameters of our 1D-CNN model.

| No. Layer | Layer type | No. kernel | Kernel size | Region size | Stride | Padding | Output size |
|-----------|------------|------------|-------------|-------------|--------|----------|-------------|
| 1         | 1D-Conv    | 64         | 1×11        | -           | 4      | 2        | 64×1×959   |
| 2         | BatchNorm  | -          | -           | -           | -      | -        | 64×1×959   |
| 3         | ReLU       | -          | -           | -           | -      | -        | 64×1×959   |
| 4         | MaxPool    | 64         | 1×3         | 2           | No     | -        | 64×1×479   |
| 5         | 1D-Conv    | 192        | 1×5         | 1           | 2      | 2        | 192×1×479  |
| 6         | ReLU       | -          | -           | -           | -      | -        | 192×1×479  |
| 7         | MaxPool    | 192        | 1×3         | 2           | No     | -        | 192×1×239  |
| 8         | 1D-Conv    | 384        | 1×3         | 1           | 1      | 1        | 384×1×239  |
| 9         | ReLU       | -          | -           | -           | -      | -        | 384×1×239  |
| 10        | 1D-Conv    | 256        | 1×3         | 1           | 1      | 1        | 256×1×239  |
| 11        | ReLU       | -          | -           | -           | -      | -        | 256×1×239  |
| 12        | 1D-Conv    | 256        | 1×3         | 1           | 1      | 1        | 256×1×239  |
| 13        | ReLU       | -          | -           | -           | -      | -        | 256×1×239  |
| 14        | MaxPool    | 256        | 1×3         | 2           | No     | -        | 256×1×119  |
| 15        | AvgPool    | 256        | -           | -           | -      | -        | 256×1×6    |
| 16        | Flatten    | -          | -           | -           | -      | -        | 1×1536     |
| 17        | Dense      | -          | 512         | -           | -      | -        | 1×512      |
| 18        | ReLU       | -          | -           | -           | -      | -        | 1×512      |
| 19        | Dense      | -          | 128         | -           | -      | -        | 1×128      |
| 20        | ReLU       | -          | -           | -           | -      | -        | 1×128      |
| 21        | Dense      | -          | 2           | -           | -      | -        | 1×2        |

VOLUME 8, 2020
3) 1D-CNN TRAINING SETTING
The weight parameters in our 1D-CNN were initialized leveraging the kaiming initializer [36], which could improve the converging speed of model. In addition, the parameter optimization was performed with Adam optimizers with an initial learning rate of 0.0001 [37]. The network was trained for 80 epochs with a batch size of 256. The cross entropy loss function was used in this study, which is defined as:

\[
E_n = -\frac{1}{N} \sum_{k=1}^{N} p_k \cdot \log(y_k) + (1 - p_k) \log(1 - y_k) \quad (8)
\]

where \(N\) is the number of the training samples, \(y_k\) and \(p_k\) denote the true label and the predicted label of the sample, respectively.

In order to prevent overfitting problems, we utilized the dropout and L2 regularization method. Dropout is a technique that units of the layer are randomly disconnected with specified probability during training [38]. L2 regularization is a technique that adds a regularization term after loss function to reduce the network complexity. The cross entropy loss function with L2 regularization term is defined as:

\[
E_{\ell_2} = E_n + \lambda \|w\|^2_2 \quad (9)
\]

where \(E_n\) is the basic cross entropy loss function. \(\lambda \|w\|^2_2\) is the L2 regularization term. \(\lambda\) and \(w\) are the penalty factor and the network parameters, respectively. In this study, the penalty factor \(\lambda\) was set to 0.0001. Moreover, we used the dropout of 0.5 after the first and second layer of the fully connected layer. Table 5 shows the hyper-parameters of our 1D-CNN.

| Parameter       | Value  |
|-----------------|--------|
| Optimizer       | Adam   |
| Learning rate   | 0.0001 |
| Loss function   | Cross entropy |
| Batch size      | 256    |
| L2 regularization| 0.0001 |
FIGURE 5. Intra-patient experiment. Performance comparison of our 1D-CNN leveraging the 4 subdatasets across accuracy (Acc), precision, recall, F1-score and k. We conducted paired two-sided t test to compare the average accuracies of our 1D-CNN leveraging ALL-DS, LSS-DS and SWS-DS with 1D-CNN leveraging REM-DS (* means p<0.05, no marking means no statistical significance).

IV. EXPERIMENTS AND RESULTS

Data preprocessing and feature extraction sections were implemented in Matlab R2018a environment on Intel i7-9700 @3.00GHz with 8 GB RAM. Deep learning experiments were conducted in Python environment leveraging Pytorch framework on NVIDIA GeForce GTX 1080 Ti. In addition, we used SPSS Software System version 20.0 for paired two-sided t test. Values of $p<0.05$ were considered statistically significant.

A. PERFORMANCE METRICS

In this research, accuracy, precision, recall, F1 score and kappa coefficient (k) were used for evaluating the performance metrics, which are defined as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$  \hspace{1cm} (10)

$$\text{Precision} = \frac{TP}{TP + FP}$$  \hspace{1cm} (11)

$$\text{Recall} = \frac{TP}{TP + FN}$$  \hspace{1cm} (12)

$$\text{F}_1\text{score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$  \hspace{1cm} (13)

$$k = \frac{p_o - p_e}{1 - p_e}$$  \hspace{1cm} (14)

where $TP$, $FP$, $TN$, $FN$ represent the true positive, false positive, true negative and false negative, respectively. $p_o$ and $p_e$ represent the actual agreement and the chance agreement, respectively.

B. EVALUATION PARADIGMS

1) INTRA-PATIENT PARADIGM

Under intra-patient paradigm, subjects are firstly split into training and testing set. The epochs from the training set and testing set subjects are used for training and testing, respectively, i.e. epochs from different patients are utilized for training and testing. Generally, inter-patient paradigm can prevent the problem of the signal similarity from the same subject, which can guarantee the generalizability of the model [7].

C. INTRA-PATIENT EXPERIMENT

1) DATASET

In intra-patient experiment, epochs from all the 18 subjects were mixed together. There were 11752 epochs in total, where insomnia patients contained 5970 samples whereas healthy subjects contained 5782 samples. 10-fold cross validation was employed to evaluate the performance of our model, i.e. each fold, in turn, was used for testing whereas the remaining 9 folds were used for training. We performed the experiment with our 1D-CNN leveraging the 4 subdatasets: ALL-DS, REM-DS, LSS-DS and SWS-DS, and compared the identification performance with baseline method.

2) RESULT

Fig. 5 depicts the identification performance metrics of our 1D-CNN leveraging the 4 subdatasets. Table 6 shows the performance comparison between baseline methods and our 1D-CNN leveraging the 4 subdatasets. Fig. 7 depicts the accuracy and loss curve for our 1D-CNN leveraging REM-DS in training process.

Fig. 5 and Table 6 show that, in intra-patient experiment, the average accuracies of our 1D-CNN leveraging REM-DS, LSS-DS and SWS-DS are 98.98%, 96.38%
B. Yang, H. Liu: Automatic Identification of Insomnia Based on Single-Channel EEG Labelled

**FIGURE 6.** Inter-patient experiment. Performance comparison of our 1D-CNN leveraging 4 subdatasets across Acc, precision, recall, F1-score and k.

**TABLE 6.** Intra-patient experiment. Performance comparison between baseline methods and our 1D-CNN leveraging the 4 subdatasets across accuracy (Acc), precision, recall, F1-score and k.

| Method                        | Subdataset | Acc(%)   | Precision | Recall   | F1-score | k        |
|-------------------------------|------------|----------|-----------|----------|----------|----------|
| Our method                    | ALL-DS     | 91.31 ± 1.12 | 0.91 ± 0.03 | 0.91 ± 0.03 | 0.92 ± 0.01 | 0.82 ± 0.02 |
|                               | REM-DS     | 98.98 ± 0.72 | 0.99 ± 0.01 | 0.99 ± 0.01 | 0.99 ± 0.01 | 0.98 ± 0.01 |
|                               | LSS-DS     | 96.38 ± 0.72 | 0.97 ± 0.01 | 0.96 ± 0.02 | 0.96 ± 0.01 | 0.93 ± 0.01 |
|                               | SWS-DS     | 99.16 ± 0.39 | 0.99 ± 0.01 | 0.99 ± 0.01 | 0.99 ± 0.01 | 0.98 ± 0.01 |
| SSA-ANN [12]                  | BSL-DS     | 84.54 ± 2.45 | 0.87 ± 0.02 | 0.82 ± 0.06 | 0.84 ± 0.03 | 0.71 ± 0.07 |
| Hjorth parameters and power   | SWS-DS     | 62.65 ± 1.51 | 0.71 ± 0.03 | 0.44 ± 0.05 | 0.54 ± 0.03 | 0.25 ± 0.02 |
| features-PCA [13]             |            |           |           |          |          |          |
| Hjorth parameters and power   | SWS-DS     | 93.02 ± 0.50 | 0.92 ± 0.01 | 0.94 ± 0.01 | 0.93 ± 0.01 | 0.86 ± 0.01 |
| features-RF [13]              |            |           |           |          |          |          |

In order to further examine any significant difference in the average accuracy of our 1D-CNN leveraging the 4 subdatasets, paired two-sided t test was conducted. Fig. 5 indicate that amongst the 4 subdatasets, the average accuracy of our 1D-CNN leveraging REM-DS is significantly higher than that of ALL-DS and LSS-DS ($p < 0.05$), whereas where is no statistically significant difference between REM-DS and SWS-DS, i.e. our 1D-CNN leveraging REM and SWS epochs exhibit the best insomnia identification performance in intra-patient experiment.

**D. INTER-PATIENT EXPERIMENT**

1) DATASET

In inter-patient experiment, subjects were firstly split into training set and testing set. Based on the leave one subject out cross validation (LOSOCV) strategy, at each time, we randomly selected one insomnia patient and one healthy subject for testing, whereas the remaining subjects were all used for training. The experiments were repeated 10 times.

We performed the experiment with our 1D-CNN leveraging the 4 subdatasets: ALL-DS, REM-DS, LSS-DS and SWS-DS, and compared the performance with baseline method.

2) RESULT

Fig. 6 depicts the identification performance metrics of our 1D-CNN leveraging the 4 subdatasets. Table 7 shows the performance comparison between baseline methods and our 1D-CNN leveraging the 4 subdatasets. Fig. 8 depicts the accuracy and loss curve for our 1D-CNN leveraging REM-DS in training process.

Fig. 6 and Table 7 show that, in inter-patient experiment, the average accuracies of our 1D-CNN leveraging ALL-DS, REM-DS, LSS-DS and SWS-DS are 86.82%, 87.49%, 82.06%, 83.76%, respectively, which are higher than that of 3 baseline methods.

Similar to intra-patient experiment, paired two-sided t test was conducted to examine any significant difference in the average of our 1D-CNN leveraging the 4 subdatasets. However, in inter-patient experiment, the average accuracies and 99.16%, respectively, which are higher than the that of 3 baseline methods.
**TABLE 7.** Inter-patient experiment. Performance comparison between baseline methods and our 1D-CNN leveraging the 4 subdatasets across accuracy (Acc), precision, recall, F1-score and k.

| Method                  | Subdataset | Acc(%)     | Precision | Recall | F1-score | k     |
|-------------------------|------------|------------|-----------|--------|----------|-------|
|                         | ALL-DS     | 86.82 ± 5.43 | 0.89 ± 0.04 | 0.81 ± 0.12 | 0.85 ± 0.08 | 0.71 ± 0.13 |
|                         | REM-DS     | 87.49 ± 4.63 | 0.90 ± 0.06 | 0.84 ± 0.10 | 0.86 ± 0.05 | 0.73 ± 0.08 |
|                         | LSS-DS     | 82.06 ± 7.33 | 0.82 ± 0.10 | 0.80 ± 0.15 | 0.80 ± 0.10 | 0.62 ± 0.15 |
|                         | SWS-DS     | 83.76 ± 11.21 | 0.82 ± 0.13 | 0.81 ± 0.18 | 0.79 ± 0.18 | 0.64 ± 0.24 |
| Our method              | BSL-DS     | 76.85 ± 19.38 | 0.72 ± 0.31 | 0.66 ± 0.35 | 0.67 ± 0.34 | 0.52 ± 0.38 |
|                         | Hjorth params and power features-PCA [13] | SWS-DS | 78.57 ± 17.54 | 0.76 ± 0.19 | 0.84 ± 0.19 | 0.78 ± 0.16 | 0.55 ± 0.35 |
|                         | Hjorth params and power features-RF [13] | SWS-DS | 73.93 ± 12.40 | 0.70 ± 0.23 | 0.66 ± 0.30 | 0.67 ± 0.26 | 0.43 ± 0.30 |

**FIGURE 7.** Intra-patient experiment. Accuracy and loss curve for our 1D-CNN leveraging REM-DS in training process.

**FIGURE 8.** Inter-patient experiment. Accuracy and loss curve for our 1D-CNN leveraging REM-DS in training process.

EEG signal labelled with sleep stage annotations. In order to further investigate which sleep stage has better insomnia identification performance, we constructed 4 subdatasets from EEG epochs according to their sleep stage annotations, and compared the performance of our 1D-CNN leveraging the 4 subdatasets respectively. Table 8 shows the comparison of identification performance for our method and other existing methods. As shows in Table 8, 1D-CNN model introduced in this paper could achieve superior performance than existing methods.

We employed two strategies when constructing subdatasets. Those two strategies guaranteed the 4 subdatasets have the same size and the same epoch number from each subject, which could avoid the effect of dataset size and difference between subjects on identification performance. In general, 1D-CNN might learn more features from the subjects who have more epochs when the epoch number from each subject varies greatly, thus leading to the performance degradation in inter-patient experiment. Therefore, we set a threshold for epoch number at 800 to prevent the great difference of the epoch number for each subject in subdataset.

Our experiments demonstrated that our 1D-CNN leveraging the 3 subdatasets composed of REM, LSS and SWS epochs, respectively, achieved higher average accuracies in comparison with baseline methods both in intra-patient and inter-patient experiments. We consider this is because 1D-CNN is an end-to-end leaning model, i.e. the feature extractor and classifier are integrated into one single algorithm. The end-to-end learning method can overcome the limitation of the handcrafted features that are limited by prior

**V. DISCUSSION**

In this paper, we proposed a 1D-CNN model for automatic insomnia identification leveraging single-channel EEG signal labelled with sleep stage annotations. In order to further investigate which sleep stage has better insomnia identification performance, we constructed 4 subdatasets from EEG epochs according to their sleep stage annotations, and compared the performance of our 1D-CNN leveraging the 4 subdatasets respectively. Table 8 shows the comparison of identification performance for our method and other existing methods. As shows in Table 8, 1D-CNN model introduced in this paper could achieve superior performance than existing methods.

We employed two strategies when constructing subdatasets. Those two strategies guaranteed the 4 subdatasets have the same size and the same epoch number from each subject, which could avoid the effect of dataset size and difference between subjects on identification performance. In general, 1D-CNN might learn more features from the subjects who have more epochs when the epoch number from each subject varies greatly, thus leading to the performance degradation in inter-patient experiment. Therefore, we set a threshold for epoch number at 800 to prevent the great difference of the epoch number for each subject in subdataset.

Our experiments demonstrated that our 1D-CNN leveraging the 3 subdatasets composed of REM, LSS and SWS epochs, respectively, achieved higher average accuracies in comparison with baseline methods both in intra-patient and inter-patient experiments. We consider this is because 1D-CNN is an end-to-end leaning model, i.e. the feature extractor and classifier are integrated into one single algorithm. The end-to-end learning method can overcome the limitation of the handcrafted features that are limited by prior
knowledge, thus improving the performance of insomnia identification.

The experimental results also indicated that amongst all the sleep stages, 1D-CNN leveraging REM and SWS epochs exhibited the best insomnia identification performance in intra-patient experiment, whereas no statistically significant difference was found in inter-patient experiment. Several researches in the past [40]–[42] demonstrated that the power of EEG during sleep between insomnia patients and healthy subjects had significant difference in NREM and REM stages. More specifically, the high frequency EEG activities (in the sigma and beta range) were increased in NREM and REM stage in insomnia group [24]. This might be a possible reason for REM-DS and SWS-DS have better insomnia identification performance in intra-patient experiment. Given the strong hereditary and individual variability, it is hard to discover statistical rules when leveraging small sample sizes dataset in inter-patient experiment. Therefore, larger sample size dataset is required for further investigation.

Moreover, the comparison of the two experiments demonstrated that the average identification accuracy of our 1D-CNN could achieve 99.16% in intra-patient experiment, whereas it could only reach 87.49% in inter-patient experiment with larger standard deviation. We consider the high accuracy in intra-patient experiment is caused by the similarity of epochs, i.e. epochs from the same patient are utilized both for training and testing. However, inter-patient experiment is more realistic evaluation paradigm which could guarantee the generalizability of the method. Therefore, we suggest future research on automatic insomnia identification based on deep learning should focus on the inter-patient experiment performance.

We want to also mention that the goal of the proposed model in this paper is to identify each 30s epoch in the whole night EEG recording correctly. If the automatic insomnia identification result is only based on one 30s epoch, it will be a waste of the practical measurement EEG data. Therefore, in the clinical practice, we suggest that each epoch of the whole night EEG recording is identified by the 1D-CNN, and then the final identification result is obtained by majority voting.

Our research also has some limitations. Firstly, the effect of the frequency band of EEG on automatic insomnia identification based on 1D-CNN has not been explored yet. Perlis et al. [42] explained the pathological mechanism of insomnia from a neurocognitive perspective, i.e. insomnia was associated with high frequency activity of EEG. In future work, we will focus on the effect of frequency band of EEG on automatic insomnia identification based on 1D-CNN. Secondly, the dataset utilized in this study was relatively small, and this is the reason why we implemented the inter-patient experiment based on LOSOCV strategy. Additionally, we tried the large-scale CNN structures such as VggNet and ResNet. However, we found that they had a good performance in intra-patient experiment, whereas they failed to produce acceptable results when it came to inter-patient experiment. We consider this is because the deep CNN structure result in overfitting over the relatively small dataset. Therefore, we reduced the scale of the network and found the 1D-CNN with 5 convolution layers could achieve superior performance. In future work, we plan to obtain larger sleep databases with sleep stage annotations. Under large sample dataset, we could select more subjects for testing, and further increase the scale of our 1D-CNN to maximize the ability of deep convolutional neural networks.

VI. CONCLUSION

In this paper, we proposed a 1D-CNN model for automatic insomnia identification based on single-channel EEG labelled with sleep stage annotations, and further investigated the identification performance based on different sleep stages. Our experiments demonstrated that our 1D-CNN leveraging the 3 subdatasets composed of REM, LSS and SWS epochs, respectively, achieved higher average accuracy in comparison with baseline methods under both intra-patient and inter-patient paradigms. The experimental results also indicated that amongst all the sleep stages, 1D-CNN leveraging REM and SWS epochs exhibited the best insomnia identification performance in intra-patient paradigm, whereas no statistically significant difference was found in inter-patient paradigm.

Overall, for automatic insomnia identification based on single-channel EEG labelled with sleep stages, 1D-CNN model introduced in this paper could achieve superior performance than traditional methods. Further experiment based on larger sleep databases under inter-patient paradigm is still required in future work.

REFERENCES

[1] E. Mignot, “Why we sleep: The temporal organization of recovery,” PLoS Biol., vol. 6, no. 4, p. e106, Apr. 2008.
[2] M. M. Ohayon, “Epidemiology of insomnia: What we know and what we still need to learn,” Sleep Med. Rev., vol. 6, no. 2, pp. 97–111, May 2002.
[3] J. D. Edinger, M. H. Bonnet, R. R. Bootzin, K. Doghranji, C. M. Dorsey, C. A. Espie, A. O. Jamieson, W. V. McCaI, C. M. Morin, and E. J. Stepanski, “Derivation of research diagnostic criteria for insomnia: Report of an American academy of sleep medicine work group,” Sleep, vol. 27, no. 8, pp. 1567–1596, Dec. 2004.
[4] V. Ihañez, J. Silva, and O. Caulti, “A survey on sleep assessment methods,” PeerJ, vol. 6, p. e4849, May 2018.
[5] I. Chouvarda, M. O. Mendez, V. Rosso, A. M. Bianchi, L. Parrino, A. Grassi, M. G. Terzano, S. Cerutti, and N. Maglaveras, “Cyclic alternating patterns in normal sleep and insomnia: Structure and content differences,” IEEE Trans. Neural Syst. Rehabil. Eng., vol. 20, no. 5, pp. 642–652, Sep. 2012.
[6] J. Newell, O. Mairese, P. Verbanck, and D. Neu, “Is a one-night stay in the lab really enough to conclude? First-night effect and night-to-night variability in polysomnographic recordings among different clinical population samples,” Psychiatry Res., vol. 200, nos. 2–3, pp. 795–801, Dec. 2012.
[7] S. Faltmann and L. Chen, “Computational sleep behavior analysis: A survey,” IEEE Access, vol. 7, pp. 142421–142440, 2019.
[8] A. R. Hassan and M. A. Haque, “An expert system for automated identification of obstructive sleep apnea from single-lead ECG using random under sampling boosting,” Neurocomputing, vol. 235, pp. 122–130, Apr. 2017.
[9] M. B. B. Heyat, D. Lai, F. I. Khan, and Y. Zhang, “Sleep bruxism detection using decision tree method by the combination of C4-P4 and C4-A1 channels of scalp EEG,” IEEE Access, vol. 7, pp. 102542–102553, 2019.
[10] S. Keenan and M. Hirshkowitz, “Monitoring and staging human sleep,” Princ. Pract. Sleep Med., vol. 10, no. 5, pp. 1602–1609, 2011.
[11] Y. Lu, M. Wang, W. Wu, Y. Han, Q. Zhang, and S. Chen, “Dynamic entropy-based pattern learning to identity emotions from EEG signals across individuals,” Measurement, vol. 150, Jan. 2020, Art. no. 107003.
[12] S. Aydin, H. M. Saraoou, and S. Kara, “Singular spectrum analysis of sleep EEG in insomnia,” J. Med. Syst. vol. 35, no. 4, p. 457, 2011.
[13] S. T.-B. Hamida, M. Glos, T. Penzel, and B. Ahmed, “How many sleep stages do we need for an efficient automatic insomnia diagnosis?” in Proc. 38th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), Aug. 2016, pp. 2431–2434.
[14] M. Shahin, B. Ahmed, S.-T. Hamida, F. L. Mulaflah, M. Glos, and T. Penzel, “Deep learning and insomnia: Assisting clinicians with their diagnosis!,” IEEE J. Biomed. Health Inform., vol. 21, no. 6, pp. 1546–1553, Nov. 2017.
[15] B.-T. Zhang, X.-P. Wang, Y. Shen, and T. Lei, “Dual-modal physiological feature fusion-based sleep recognition using CFS and RF algorithm,” Int. J. Autom. Comput., vol. 16, no. 3, pp. 286–296, Jun. 2019.
[16] A. Craik, Y. He, and J. L. Contreras-Vidal, “Deep learning for electroencephalogram (EEG) classification tasks: A review,” J. Neural Eng., vol. 16, no. 3, Jun. 2019, Art. no. 031001.
[17] R. Zemouri, N. Zerhouni, and D. Racoceanu, “Deep learning in the biomedical applications: Recent and future status,” Appl. Sci., vol. 9, no. 8, p. 1526, Apr. 2019.
[18] W. Gu, L. Shangguan, Z. Yang, and Y. Liu, “Sleep hunter: Towards fine grained sleep stage tracking with smartphones,” IEEE Trans. Mobile Comput., vol. 15, no. 6, pp. 1514–1527, Jun. 2016.
[19] J. Allan Hobson, “A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects,” Electroencephalogr. Clin. Neurophysiol., vol. 26, no. 6, p. 644, Jun. 1969.
[20] A. R. Hassan and M. I. H. Bhuiyan, “Automated identification of sleep states from EEG signals by means of ensemble empirical mode decomposition and random under sampling boosting,” Comput. Methods Programs Biomed., vol. 140, pp. 201–210, Mar. 2017.
[21] D. Jiang, Y.-N. Lu, Y. Ma, and Y. Wang, “Robust sleep stage classification with single-channel EEG signals using multimodal decomposition and HMM-based refinement,” Expert Syst. Appl., vol. 121, pp. 188–203, May 2019.
[22] A. Supratak, H. Dong, C. Wu, and Y. Guo, “DeepSleepNet: A model for automatic sleep stage scoring based on raw single-channel EEG,” IEEE Trans. Neural Syst. Rehabil. Eng., vol. 25, no. 11, pp. 1998–2008, Nov. 2017.
[23] K. Chen, C. Zhang, J. Ma, G. Wang, and J. Zhang, “Sleep staging from single-channel EEG with multi-scale feature and contextual information,” Sleep Breathing, vol. 23, no. 4, pp. 1159–1167, Dec. 2019.
[24] B. Feige, C. Baglioni, K. Spiegelhalter, V. Hirscher, C. Nissen, and D. Racoceanu, “M.S. degree in electronics and communication engineering with Nanjing University, China. He is also a member of the Biomedical Electronics Branch, Chinese Institute of Electronics. He has focused on fetal electrocardiogram technology and achieved remarkable results. He has presided over three projects of the National Natural Science Foundation, one social development project of Jiangsu Science and Technology Support Program, and more than ten horizontal cooperation projects. He has published nearly 100 articles by the first author or correspondent author, including more than 20 papers by SCI. The first inventor applied for 33 invention patents, of which 17 were authorized. His research interests include intelligent information processing and biomedical electronics.

BUFANG YANG received the B.S. degree in detection, guidance and control technology from Northwestern Polytechnical University, China, in 2019. He is currently pursuing the M.S. degree in electronics and communication engineering with Nanjing University, China. His current research interests include electroencephalogram signal processing, machine learning, deep learning, and automatic sleep stage scoring.

HONGXING LIU received the Ph.D. degree from Xi’an Jiaotong University, China, in 1997. He is currently a Professor of information electronics with the School of Electronic Science and Engineering, Nanjing University, China. He is also a master’s tutor for signal and information processing. He is also a member of the Biomedical Electronics Branch, Chinese Institute of Electronics. He has focused on fetal electrocardiogram technology and achieved remarkable results. He has presided over three projects of the National Natural Science Foundation, one social development project of Jiangsu Science and Technology Support Program, and more than ten horizontal cooperation projects. He has published nearly 100 articles by the first author or correspondent author, including more than 20 papers by SCI. The first inventor applied for 33 invention patents, of which 17 were authorized. His research interests include intelligent information processing and biomedical electronics.

VOLUME8, 2020