IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 31, 2019, accepted June 20, 2019, date of publication July 11, 2019, date of current version August 5, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2928129

Pediatric Sleep Stage Classification Using Multi-Domain Hybrid Neural Networks

YONGHOON JEON[1], SIWON KIM[2], HYUN-SOO CHOI[2], YOON GI CHUNG[1], SUN AH CHOI[3,4], HUNMIN KIM[3], SUNGROH YOON[2,5], (Senior Member, IEEE), HEE HWANG[3], AND KI JOONG KIM[6,7]

[1]Healthcare ICT Research Center, Seoul National University Bundang Hospital, Seongnam 13605, South Korea
[2]Department of Electrical and Computer Engineering, Seoul National University, Seoul 08826, South Korea
[3]Department of Pediatrics, Seoul National University Bundang Hospital, Seongnam 13620, South Korea
[4]Department of Pediatrics, Dankook University Hospital, Cheonan 31116, South Korea
[5]ASRI, INMC, ISRC, and Institute of Engineering Research, Seoul National University, Seoul 08826, South Korea
[6]Pediatric Clinical Neuroscience Center, Seoul National University Children's Hospital, Seoul 03080, South Korea
[7]Department of Pediatrics, Seoul National University College of Medicine, Seoul 03080, South Korea

Corresponding authors: Sungroh Yoon (sryoon@snu.ac.kr) and Hee Hwang (neuroandy@snubh.org)

This work was supported in part by the Institute for Information and Communications Technology Promotion (IITP) Grant funded by the Korean Government [Ministry of Science and ICT (MSIT)] (Intelligent SW Technology Development for Medical Data Analysis) under Grant 2018-0-00861, in part by the Brain Korea 21 Plus Project in 2019, in part by the National Research Foundation of Korea (NRF) Grant funded by the Korean Government (MSIT) under Grant 2018R1A2B3001628, in part by the Projects for Research and Development of Police Science and Technology through the Center for Research and Development of Police Science and Technology and the Korean National Police Agency funded by the Ministry of Science, ICT and Future Planning under Grant PA-C000001, and in part by the Hyundai Motor Company.

ABSTRACT Sleep staging is an important part of clinical neurology. However, it is still performed manually by technical experts and is labor-intensive and time-consuming. To overcome these obstacles in the manual sleep staging process, a large number of machine learning-based classifiers with hand-engineered features have been proposed. Additionally, combinations of a deep neural network (DNN) have been recently highlighted as the state-of-the-art classifiers in view of their effectiveness for automatic sleep staging. In spite of the existence of a large number of these types of classifiers, to-this-date, no prior DNN-based approach has attempted sleep-stage classification using pediatric electroencephalographic (EEG) signals. In this paper, we propose a novel end-to-end classifier based on a multi-domain hybrid neural network (HNN-multi) approach consisting of a convolutional neural network and bidirectional long short-term memory for automatic sleep staging with pediatric scalp EEG recordings. To find effective temporal, spatial, and domain-specific conditions, we investigated noticeable changes in the classification performance corresponding to: 1) the length of input signals; 2) the number of channels; and 3) the types of input signals in the time and frequency domains. Our HNN-based classifier yielded the best performance metrics using 30-s time series in combination with an instantaneous frequency using a 19-channel, three-stage classification, with overall accuracy, F1 score, and Cohen's Kappa, equal to 92.21%, 0.90, and 0.88, respectively. We suggest that an effective combination of temporal and spatial time-domain clues with time-varying frequency domain information plays a pivotal role in pediatric, automatic sleep staging. Sufficiently reasonable performance of our HNN-based approach coping with the highly complicated pediatric EEG signatures hopefully sheds light on the clinical feasibility of the DNN-based automatic sleep staging for pediatric neurology.

INDEX TERMS Automatic sleep staging, deep learning, convolutional neural network, long short-term memory, instantaneous frequency features, pediatric electroencephalography.

I. INTRODUCTION

Sleep staging in electroencephalography (EEG) is an important initial step for the identification of vigilance and subsequent analyses in the field of clinical neurology. Wakefulness and sleep stages exhibit a broad range of distinct electroencephalographic features [1]–[3]. Accurate identification of each sleep stage plays a critical role in clinical interpretations and neurological studies. However, current sleep staging is performed manually by technicians in

The associate editor coordinating the review of this manuscript and approving it for publication was Yongqiang Cheng.

 This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see http://creativecommons.org/licenses/by/4.0/

polysomnography (PSG), and is labor-intensive and time-consuming. Although the process of sleep staging follows strict criteria proposed by the American Academy of Sleep Medicine (AASM) [4], it is likely that inaccurate sleep staging outcomes may be evoked. In addition, electroencephalographers and neurologists perform the manual sleep staging with other EEG interpretation steps simultaneously, such as artifact rejection and feature recognition, for in-depth EEG analysis, which may hamper their efficiency by burdensome manual tasks. These potential risks, errors, and laborious characteristics of the manual sleep staging have urgently raised the need for automatic sleep staging.

Previous studies in automatic sleep staging generally consisted of feature extraction followed by machine learning-based classification. Prominent features for each sleep stage have been explored in the time, frequency, and time–frequency domains [5]–[15] and achieved accuracy > 87% in sleep stage classification procedures in healthy subjects based on various classical machine learning-based classifiers [16]. Thanks to the emergence and advents of deep-learning technology, state-of-the-art deep neural network (DNN) have been adopted in recent studies [17]–[28] compared to classical machine learning-based classifiers either with hand-engineered features, or in an end-to-end manner with self-generated features [29].

To successfully utilize the temporal characteristics of EEG time series, a hybrid neural network (HNN) combined with convolutional neural network (CNN) and recurrent neural network (RNN)–particularly with the bidirectional long short-term memory (biLSTM) to simultaneously reflect past and future sequences–has been extensively adopted in many recent automatic sleep staging strategies. State-of-the-art HNN-based approaches have demonstrated the effectiveness of the HNN for automatic sleep staging in terms of the temporal dynamics of sleep-stage transitions. However, to-this-date, no HNN-based study has attempted automatic sleep staging with pediatric scalp EEG recordings. Ansari *et al.* [30] implemented a CNN-based classifier using EEG signals from preterm infants, but their approach was limited to a binary classification between quiet and nonquiet sleep stages.

In this study, we propose a novel HNN-based classifier, based on CNN and biLSTM, for automatic sleep staging with pediatric scalp EEG recordings. The contributions of our work are as follows:

- To the best of our knowledge, this is the first study that classifies multiclass sleep stages with pediatric scalp EEG signals which have intrinsic dynamic heterogeneity owing to brain maturation in comparison with adult signals
- We propose a novel multi-domain HNN to effectively learn the morphological and frequency information in EEG. Furthermore, we adopt instantaneous frequency (IF) to extract detailed frequency information for short-time signal segments that outperform conventional power spectral density (PSD)
- To find effective temporal, spatial, and domain-specific conditions for favorable sleep staging, we investigate the classification performance by changing the lengths of the input signals and number of input channels

II. BACKGROUND

A. TRADITIONAL MACHINE LEARNING-BASED CLASSIFIERS FOR SLEEP STAGING

Seifpour *et al.* [5] extracted time domain features, referred to as statistical behaviors of local extrema, and represented the hidden dynamics of time series from six frequency subbands of EEG recordings for six-stage classifications based on support vector machines (SVM) with an accuracy > 90%. Diykh and Li [6] generated graph-based complex networks converted from time domain statistical features for six-stage classification based on a k-means classifier that achieved accuracies > 92%. Hassan and Bhuiyan [7] computed statistical time domain features from intrinsic mode functions with the use of the ensemble, empirical mode decomposition for their proposed classifier, commonly known as random undersampling boosting, and achieved an accuracy > 88% in six-stage classification. Chriskos *et al.* [8] estimated functional connectivity including the synchronization likelihood and relative wavelet entropy in the time and frequency domains, respectively, for four-stage classification based on SVMs and k-nearest neighbors (kNNs), and achieved an accuracy > 90%. Lajnef *et al.* [9] extracted statistical measures, PSD estimates, and nonlinear features, including the permutation entropy and Teager energy operator in the time and frequency domains for five-stage classification based on the SVM with 92% accuracy. Jiang *et al.* [10] obtained statistical and nonlinear features from signal decomposition in the time and frequency domains with the empirical mode decomposition and bandpass filtering for five-stage classification based on the random forest (RF) approach followed by the hidden Markov model (HMM) with an accuracy > 89%. Memar and Faradji [11] extracted time domain statistical measures, nonlinear features, including the fractal dimension and entropy, and instantaneous phases from eight frequency subbands of EEG recordings and achieve an accuracies > 87% in six-stage classification based on the RF. Karimzadeh *et al.* [12] calculated the Shannon entropy from the IF and envelopes of four frequency subbands of EEG recordings for six-stage classification using the kNN classifiers and achieved accuracies > 88%. Bajaj and Pachori [13] implemented time-frequency images with the use of the smoothed pseudo Wigner-Ville distribution from EEG recordings for six-stage classification based on SVMs and achieved an accuracy > 88%. Fraiwan *et al.* [14] calculated Reny's entropy measures from seven frequency subbands of time-frequency distributions by the Choi-Williams distribution, continuous wavelet transform, and Hilbert-Huang transform for five-stage classifications based on the RF with an 83% accuracy. Dimitriadis *et al.* [15] proposed a phase-amplitude coupling, one of the cross-frequency coupling methods, for six-stage classification based

on a multiclass naive Bayes classifier, and achieved accuracies > 94%.

B. DEEP-LEARNING-BASED CLASSIFIERS FOR SLEEP STAGING

Malafeev *et al.* [17] showed that in DNN-based classifiers in five-stage classifications, the feature-based long short-term memory (LSTM) and raw-data-based CNN outperformed the feature-based classical classifier that consisted of the RF and HMM. Patanaik *et al.* [18] obtained a compressed feature set based on the use of spectrograms of EEG recordings as inputs to the CNN in five-stage classification, and achieved accuracies > 89%. Zhang *et al.* [19] deployed the unsupervised complex-valued feature learning based on the use of the k-means clustering algorithm, and achieved an accuracy of 87% in five-stage classification using CNN. Chriskos *et al.* [20] generated multidimensional functional connectivity images as inputs to the CNN in the default mode network in four-stage classification and achieved accuracies > 99%. Sors *et al.* [21] implemented the CNN on a single-channel time series with no preprocessing for five-stage classification, and achieved an accuracy of 87%. Chambon *et al.* [22] performed five-stage classification sleep staging based on the CNN with multivariate time series, and showed that: a) their proposed approach outperformed the conventional feature-based method and b) changes in the spatial and temporal contexts affected the classification performances. Dong *et al.* [23] implemented a mixed neural network in a five-stage classification scheme that consisted of the rectifier neural network, based on the use of the multilayer perceptron, LSTM with spectral power statistics, and Shannon's entropy in the frequency domain, and achieved accuracies > 85%. Michielli *et al.* [24] proposed a cascaded RNN architecture in five-stage classification based on the LSTM with time and frequency domain feature selections, and achieved accuracies > 86%. Supratak *et al.* [25] adopted CNN in a five-stage classification scheme for automatic, time-invariant feature extraction, in conjunction with the bidirectional LSTM, and achieved accuracies > 86%. Bresch *et al.* [26] implemented a network which consisted of CNN and LSTM and achieved a Cohen's Kappa value of 0.73 in a five-stage classification scheme. Biswal *et al.* [27] combined CNN and the RNN with the spectrogram of EEG recordings in five-stage classification and achieved an accuracy > 87%. Mousavi *et al.* [28] proposed a CNN-LSTM-based method, referred to as SleepEEGNet, and achieved accuracies > 84% in a five-stage classification scheme.

III. METHODS

A. DATA ACQUISITION

This study was approved by the Institutional Review Board (IRB) of the Seoul National University of Bundang Hospital (SNUBH) (IRB number B-1904-537-104), and conducted in accordance with the Declaration of Helsinki 2013 and International Conference on Harmonization for Good Clinical Practice (ICH-GCP). The IRB waived the need to obtain informed consent due to the retrospective nature of the review of medical records and EEG data. The dataset in this study consisted of routine scalp EEG recordings of 218 pediatric subjects in the SNUBH (115 females with a mean age ± standard deviation (sd) of 12.73±2.69, and 103 males with a mean age ± sd of 13.37±2.53). All of the EEG recordings were clinically interpreted as normal and they were free of epileptiform discharges. All the signals were recorded with a 32-channel digital EEG system (Grass Telefactor Inc., West Warwick, Rhode Island, United States) for at least 30 min using a sampling frequency of 200 Hz, a 60 Hz notch filter, and 19 electrodes, according to the international 10-20 system. The signals were re-referenced to an average reference montage for subsequent analyses.

TABLE 1. Details of number of sleep stages in each epoch length of in-house dataset.

Epoch (s)	W	N1	N2	Total
30	5,504	2,020	4,352	11,876
20	8,207	3,030	6,535	17,772
10	16,512	6,060	13,062	35,634

W: Wakefulness, N1: NREM 1, N2: NREM 2

B. EPOCHING

According to the AASM standard, two EEG board certified neurologists (HK and SC) independently labeled sleep scores based on 30 s epochs in each EEG recording. Both manual annotations yielded Kappa scores which were equal to 0.85. A third EEG board certified neurologist (HH) confirmed the final scoring. EEG data were epoched into 30 s, 20 s, and 10 s epochs to follow the AASM standard, to comply with conventional clinical practices for visual inspection, and to examine how a shorter epoch length would affect sleep stage classification, respectively. No artifact rejection was considered to avoid any possible signal distortion. According to the AASM rules, stage W (wakefulness), N1 (NREM 1), N2 (NREM 2), N3 (NREM 3), and REM (rapid eye movement) are characterized by: the presence of alpha and beta waves more than 50% in the epoch, frequent eye movements, and muscle activity; the disappearance of more than 50% of alpha waves in the epoch, the decrease in eye movements, and the appearance of vertex waves; the presence of sleep spindles and K-complexes; the dominance of high amplitude delta waves; and the presence of saw-tooth waves with rapid eye movements, respectively [16], [24]. The labeled sleep stages included stage W, stage N1, and stage N2. However, N3 and REM stages were not labeled by reason of their limited availability in our dataset. Details of the number of epochs for each sleep stage based on these three epoch criteria are shown in the Table 1.

C. CHANNEL SELECTION

To investigate the influence of the multivariate spatial information on the performance for automatic sleep staging,

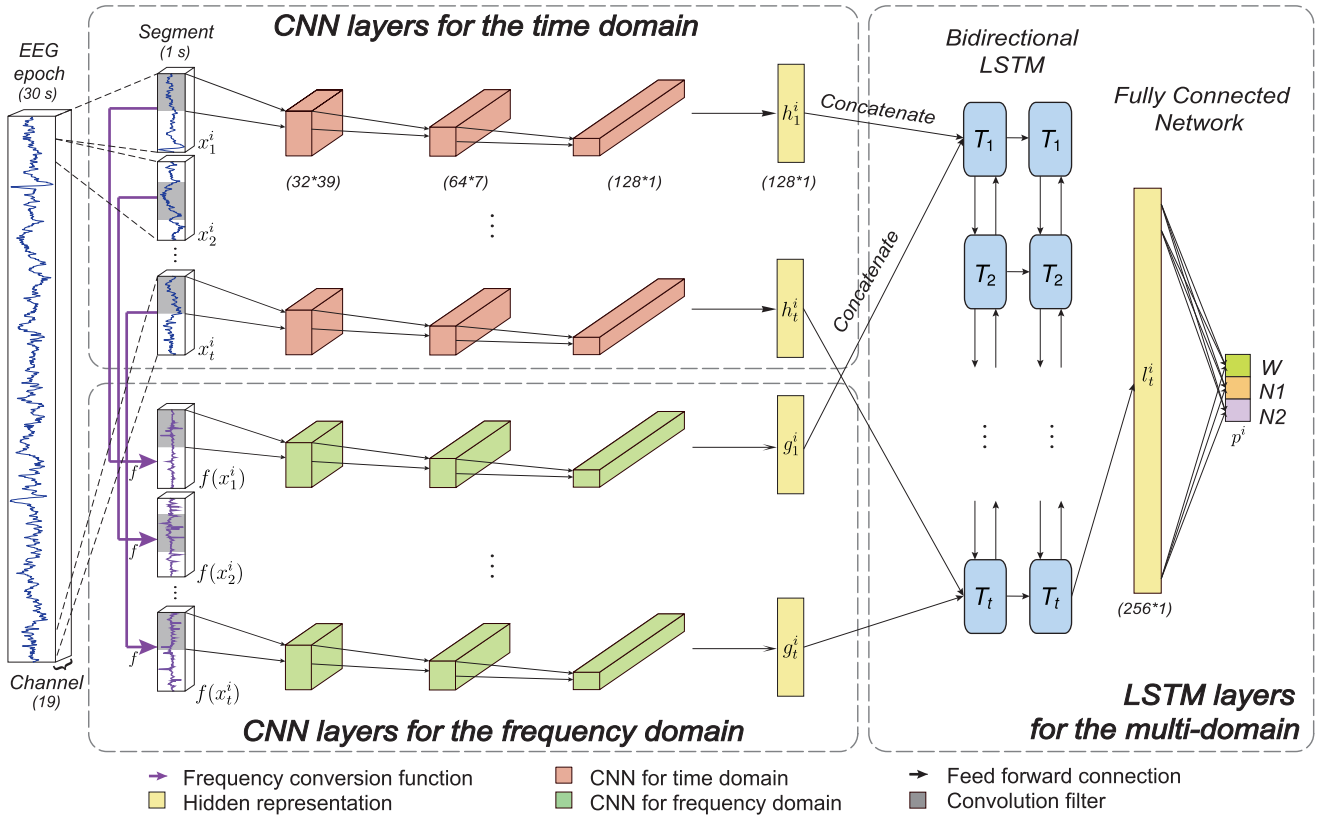


FIGURE 1. Overall Architecture.

we utilized two types of EEG channel sets as follows: (1) 19 channels including Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T3, T4, T5, T6, Fz, Cz, and Pz for the entire set of EEG channels in our dataset, (2) six channels including Fp1, Fp2, C3, C4, O1, and O2, based on the consideration of the EEG channels in the PSG. Two prefrontal EEG channels, Fp1 and Fp2, were additionally included in comparison with the PSG channels. To identify the distinct influences of scalp EEG recordings on performance, no other electrophysiological recordings, such as electrooculogram (EOG), electromyogram (EMG), and electrocardiogram (EKG), were considered in this study.

D. NETWORK ARCHITECTURES

To learn temporal dependencies as well as local time-invariant features in EEG, we introduce CNN and RNN hybrid network. Furthermore, as the EEG is a signal where not only raw signal itself but also its frequency contents are important, we design a multi-domain HNN, namely HNN-multi. It consists of two independent CNNs for modeling time and frequency domain information, respectively at the same time. The overall architecture can be classified into three parts: CNN layers for the time domain, CNN layers for the frequency domain, and LSTM layers for the multi-domain as depicted in Figure 1. The details of each of these three parts are described in the following subsections.

1) CNN LAYERS FOR THE TIME DOMAIN

The network architecture for modeling time domain signal is depicted in the top left part of Figure 1. To learn both the spatial features and temporal properties, we divide one epoch of an EEG signal into 1 s segments. Each segment is processed through the three layers. Each of these layers is a sequential combination of one-dimensional (1D) convolution layers with filters with size of 5, batch normalization, Rectified Linear Unit (ReLU) activation and maxpooling with the same kernel size as that of the convolution filter. The hidden representation of the time domain signal after the CNN can be written as,

$$h_j^i = \mathrm{CNN}_{\theta_t}(x_j^i) \tag{1}$$

where $x_j^i \in \mathbb{R}^{C\times L}$ denotes the j^{th} segment in the i^{th} epoch of the EEG and CNN_{θ_t} denotes the CNN used to model the raw signal which consists of three layers with parameters θ_t. C and L indicate the number of channels and the segment length, respectively.

2) CNN LAYERS FOR THE FREQUENCY DOMAIN

We converted all the segments of the raw signal into frequency information and employed an identical CNN architecture to CNN layers for the time domain, but with independent parameters. If we define the function $f(x_j^i)$ which converts the raw signal (x_j^i) into frequency information, the hidden

representation g_j^i after the CNN layers for the frequency domain is,

$$g_j^i = \mathrm{CNN}_{\theta_f}(f(x_j^i)) \tag{2}$$

where CNN_{θ_f} denotes the CNN with parameters θ_f learned from frequency information.

For the frequency conversion function $f(x)$, we employ the IF rather than the conventional PSD estimation. IF is well known to be appropriate for the consideration of the time-varying phase for signal segments with short-time spans [31]. Given that each EEG segment x_j^i is dependent on time, the analytical form $z(t)$ can be expressed as,

$$z_j^i(t) = x_j^i(t) + jH(x_j^i(t)) \tag{3}$$

where $H(x(t))$ is the Hilbert transform, i.e., the convolution of the input function with the function $1/(\pi t)$. Consequently, (3) can be written as

$$z_j^i(t) = A_j^i(t)\exp(j\phi_j^i(t)) \tag{4}$$

where $A_j^i(t)$ is the envelope and $\phi_j^i(t)$ is the phase of $z_j^i(t)$, i.e. $\arctan(H(x_j^i(t))/x_j^i(t))$, respectively. Finally, IF can be obtained by taking the scaled derivative of $\phi_j^i(t)$ with respect to time as,

$$f(x_j^i(t)) = \frac{1}{2\pi}\frac{d\phi_j^i(t)}{dt} \tag{5}$$

The network used to model frequency information is depicted in the bottom left part of Figure 1. Note that we omit the notation t for simplicity.

3) LSTM LAYERS FOR THE MULTI-DOMAIN

The final part of HNN-multi is depicted in the rightmost part of Figure 1. Finally, two hidden vectors, a modeled raw signal and frequency information are concatenated and enter an LSTM network which is designed to learn the long-term dependency within an entire epoch. We apply two layers of bidirectional LSTM which utilize not only the forward states, but also the backward states. A fully connected (FC) layer is added to classify the last time step vector into three classes, W, N1, and N2. The overall flow is as follows,

$$c_j^i = \{h_j^i; g_j^i\} \tag{6}$$

$$l^i = \mathrm{LSTM}(c^i) \tag{7}$$

$$p^i = \mathrm{FC}(l_t^i) \tag{8}$$

where $c^i = \{c_1^i, c_2^i, \ldots, c_N^i\}$ denotes the concatenated vector of h^i and g^i. The vector from the last time step of LSTM, l_t^i, enters the FC layer with a softmax activation that encodes the output to a probability that is used to calculate the loss to update the network parameters.

E. PERFORMANCE EVALUATION METRICS

To evaluate the performance of the proposed architecture, we used three metrics: the overall accuracy, F1 score, and Cohen's Kappa. F1 score is a harmonic mean of the precision and recall as formulated below,

$$\text{F1 score} = \frac{2 \cdot \text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \tag{9}$$

which can consider data imbalances between classes [32]. Given that we have formulated the problem as a multi-class classification problem, we calculate the accuracy and F1 score for each class and use the class mean values. Cohen's Kappa (κ) measures the agreement between two observers, considering the possibility of the agreement occurring coincidentally [33]. In our case, we measure the agreement between labels generated by neurologists and the ones predicted by the proposed HNN-multi. Accordingly, κ can be calculated as follows,

$$\kappa = \frac{p_0 - p_e}{1 - p_e} \tag{10}$$

where p_o is identical to accuracy and p_e denotes the probability of accidental agreement. If there is perfect agreement, κ becomes equal to unity.

IV. RESULTS

A. IMPLEMENTATION DETAILS

Our in-house dataset had an imbalanced distribution of sleep stages. In particular, the number of N1 stages was much smaller than those of the other stages. To handle the issue of the biased performance on machine learning-based models that resulted from the imbalanced dataset, we split our dataset based on a stratified sampling method [34]. All the epochs in our dataset were randomly split into training, validation, and test sets at a ratio of 7:1:2. Our HNN-based model was implemented in Python with Pytorch [35] based on a deep learning framework, which was trained with a maximum of 500 epochs. An AMSGrad [36] optimizer was used to minimize the categorical cross-entropy loss with mini-batches with a size 128. In this study, we ran the model on a machine with 10 central processing units (CPUs) (Intelő Xeonő CPU @ 2.20GHz), on an Ubuntu 16.04 platform. We also used four graphic processing units (GPUs) (NVIDIA TITAN Xp) to accelerate the processing of the experiments.

B. INFLUENCES OF THE INPUT COMBINATION ON MODEL PERFORMANCE

We conducted evaluations on different input combinations for HNN. The results are shown in Table 2. We evaluated and compared the power spectral density (PSD) additionally to compare with the effects of IF. PSD was calculated with the use of Welch's method [37] and with a Hamming window for a 1 s segment with a 50% overlap. When a single type of input was used (corresponding to rows 1–3 of Table 2), we applied HNN with only one CNN. For example, when time was the only input used, frequency modeling CNN was eliminated and vice versa. Note that the performance was measured with test data from the identical data split.

As shown in Table 2, using only the PSD as an input yielded the worst performance compared to those evoked

TABLE 2. Performance comparison with various input combination.

Input	Accuracy	F1 score	Kappa
Time	0.9196	0.8898	0.8705
PSD	0.6006	0.4614	0.4182
IF	0.8985	0.8617	0.8364
Time + PSD	0.8729	0.8123	0.7935
Time + IF	**0.9221**	**0.8954**	**0.8712**

Time: EEG Raw Signal, PSD: Power Spectral Density, IF: Instantaneous Frequency

from the various other inputs. This indicates that the PSD calculated with short-time segment resulted in a loss of temporal information compared to the time domain signal (Time) and IF. Owing to the low expressive power of the PSD, it yielded a worse performance when it was used together with Time in multi-domain HNN compared to the case at which only the raw signal was used. By contrast, IF showed a reasonable performance when it was used alone which led to the best performance when it was used in combination with the raw signal as an input to the HNN-multi.

C. INFLUENCES OF THE INPUT LENGTH, NUMBER OF CHANNELS, AND FREQUENCY DOMAIN INFORMATION

Table 4 shows the performance of our HNN model for three-stage classification according to the three studied metrics, i.e., the overall accuracy, F1 macro, and kappa, according to the variation of the input lengths, number of channels, and the types of input signals. Input signals combined with the 19-channel time series and IF achieved the best classification performance. The overall accuracy of the 19-channel input signals with time series, IF, and time series in combination with IF decreased from 91.96% to 91.19%, 89.86% to 88.64%, and from 92.21% to 90.23%, respectively, according to the changes of the input lengths from 30 s to 10 s. The overall accuracy of the six-channel input signals with time series, IF, time series in combination with IF increased when the input lengths decreased from 30 s to 20 s, and then decreased when the input lengths decreased from 20 s to 10 s. For all the types of input signals, 19-channel input signals guaranteed the best performance in the input length of 30 s, whereas six-channel input signals guaranteed the best performance at an input length of 20 s. Regarding to the types of input signals, the time series in combination with IF guaranteed the best performance with 19-channel input signals for an input length of 30 s, whereas the time series guaranteed the best performance with six-channel input signals for an input length of 20 s. For a single channel, Fp2 showed the best performance with 88.85%, 0.8571 and 0.82 of the overall accuracy, F1 macro, kappa, respectively using 30 s multi-domain HNN.

Confusion matrices of our HNN model for three-stage classification are shown in Figure 2 including the number of epochs and performance rates. The highest performance for waking stage (0.97) was maintained according to the changes in the types of input signals, whereas the performance for the N1 stage increased from 0.76 to 0.81 based on the use of the time series in combination with IF for 19-channel input signals. The performance for the waking stage slightly increased from 0.96 to 0.97, whereas the performance of the N1 stage decreased from 0.82 to 0.75 based on the use of the time series in combination with IF in the case of the six-channel input signals. Given that there was no remarkable variation in the performance for waking and N2 stages according to the changes in the length of the input signals and the number of channels, the enhancement of the overall accuracy depended considerably on the performance of the N1 stage.

TABLE 3. HNN model performance comparison.

Input length (s)	Channel	Score	Time	IF	Time + IF
30	19	Accuracy	0.9196	0.8986	**0.9221**
		F1 score	0.8898	0.8617	**0.8953**
		Kappa	0.8705	0.8364	**0.8752**
	6	Accuracy	0.9083	0.8834	0.9057
		F1 score	0.8789	0.8419	0.8747
		Kappa	0.8537	0.8123	0.8490
20	19	Accuracy	0.9125	0.8835	0.9111
		F1 score	0.8819	0.8354	0.8810
		Kappa	0.8598	0.8111	0.8576
	6	Accuracy	0.9159	0.8847	0.9128
		F1 score	0.8888	0.8426	0.8818
		Kappa	0.8657	0.8143	0.8595
10	19	Accuracy	0.9119	0.8864	0.9023
		F1 score	0.8874	0.8419	0.8646
		Kappa	0.8587	0.8164	0.8427
	6	Accuracy	0.9091	0.8838	0.8990
		F1 score	0.8767	0.8409	0.8673
		Kappa	0.8540	0.8128	0.8391

Time: EEG Raw Signal, PSD: Power Spectral Density, IF: Instantaneous Frequency

TABLE 4. Performance comparison with conventional machine learning models.

	Model	Accuracy	F1 score	Kappa
Time	SVM	0.4667	0.2212	0.0076
	RF	0.7197	0.5699	0.5205
	GB	0.7618	0.6441	0.5969
	CNN	0.9133	0.8806	0.8605
	HNN	0.9196	0.8898	0.8705
Time + IF	SVM	0.4667	0.2212	0.0076
	RF	0.6321	0.5303	0.3923
	GB	0.6969	0.6058	0.5034
	CNN	0.9146	0.8832	0.8625
	HNN	**0.9221**	**0.8954**	**0.8712**
BP	SVM	0.4634	0.2111	0.0067
	RF	0.9289	0.9017	0.8851
	GB	0.9292	0.9002	0.8855

SVM: Support Vector Machine, RF: Random Forest, GB: Gradient Boost, BP: Band Power

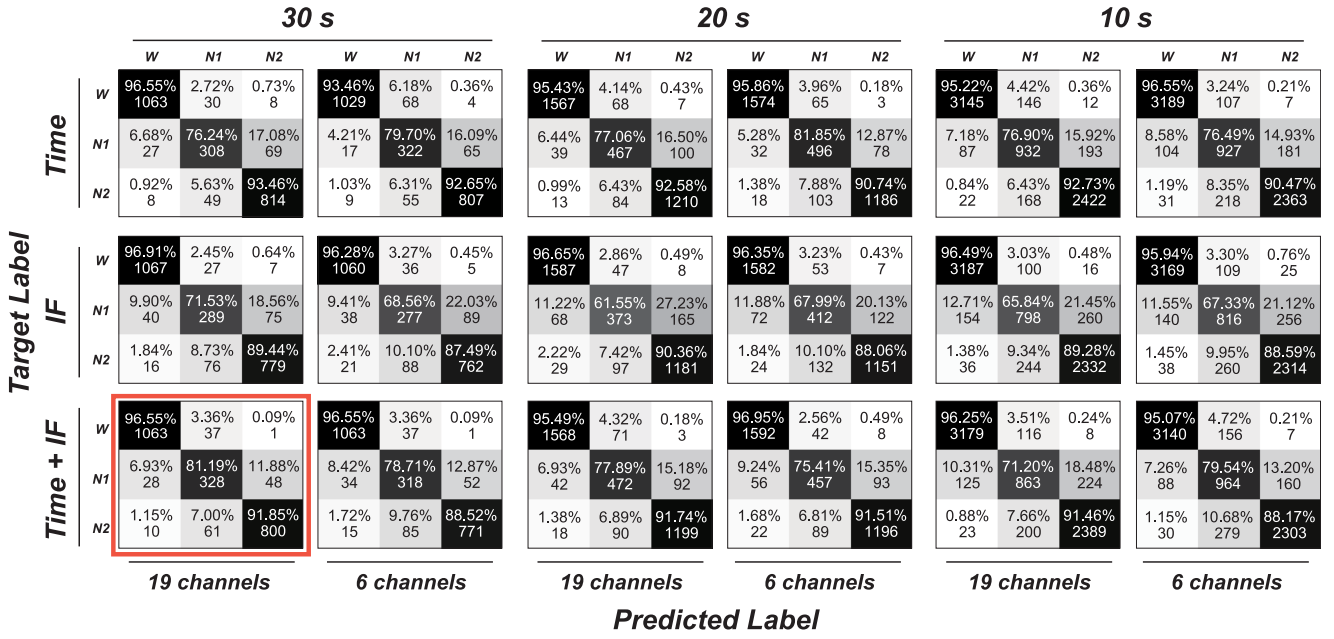


FIGURE 2. Confusion matrices of variations on epoch length and channel. The listed numbers at the bottom of each cell is the number of samples predicted to each class and the top number indicates its ratio to the sample number of the target class. The red box denotes the confusion matrix of the proposed HNN-multi which yielded the best performance.

D. PERFORMANCE COMPARISON TO DIFFERENT MODELS

An extensive search for channels and input length combinations indicated that HNN-multi with the 30 s epoch EEG from 19 channels yielded the best performance. To compare the performance to that of conventional machine learning models, we trained random forest (RF) [38], gradient boost (GB) [39], and support vector machine (SVM) [40] with data including Time and combination of Time and IF. Given that conventional machine learning models are effective when the inputs are extracted features, we decomposed the EEG into five non-overlapping frequency ranges including delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (30–50 Hz), according to a conventional approach [41]. We extracted the power of each band in each of the 19 channels, namely band power (BP), and obtained a final input vector with a dimension of 90 by concatenating these values. Additionally, to verify the effect of RNN of proposed multi-HNN, we also experimented with only CNN after removing the LSTM layers in the proposed network.

As a result, 13 combinations were tested in total, applying five models to two input types (Time and Time+IF) and using extracted BP as the input to three shallow models. HNN and CNN was not applied to BP, as deep learning is expected to automatically extract features from raw data during training. We trained each model with the identical data split. In the shallow models, we searched for the five random configurations of the hyperparameters. The results are shown in Table 4, whereby the text highlighted in boldface indicates the best performance among classifiers based on the use of raw data without feature extraction. The underlined terms indicate the best performances among the classifiers based on the use of extracted features. Conventional machine learning models yielded inferior performances in the cases of raw signal and combinations of time and IF compared to HNN. Using a combination of Time and IF as an input to HNN-multi yielded the best performance among classifiers without feature extraction. Shallow learning models, especially RF and GB seemed to show great performances with BP as highlighted with underlines in the table. However GB yielded a worse performance in N1 classification and both models have a limitation associated with saturating performance compared with HNN-multi. These outcomes are discussed in section V below. Using only CNN showed inferior performance to HNN indicating that the RNN is effective to model the temporal dependencies within the data.

V. DISCUSSION

Our HNN-based automatic sleep staging approach for pediatric scalp EEG recordings achieved the best performance using 19-channel time series input signals with a length of 30 s in combination with IF in a three-stage classification scheme. The six-channel time series signals with periods of 20 s yielded better performances compared to other conditions. In the cases of the 19-channel time series signals with periods of 30 s which were combined with IF yielded better performances compared to other conditions. According to our results, it has been considered that the performance enhancement was mainly attributed to the reduction of the error rate for N1-stage classification irrespective of the types of input signals used.

A. END-TO-END CLASSIFIER

It has been shown that many recent studies deployed state-of-the-art DNN-based approaches in an end-to-end manner

for automatic sleep staging with no hand-engineered features [17], [21], [22], [25], [26], [28]. These studies used publicly available PSG datasets with single- or multi-channel time series which were input into the CNN-based or CNN-RNN-based classifiers. Combinations of the CNN and RNN, referred to as the HNN, have been mainly adopted for the end-to-end sleep staging classification to successfully reflect temporal stage transitions [17], [22], [25], [26], [28]. Our HNN-based approach achieved an overall accuracy > 90%, which was nearly equivalent to, or outperformed the accuracies of the recently proposed methods. We consider that our performance is reasonable in terms of its potential for clinical feasibility because we used pediatric scalp EEG recordings from outpatients and not the PSG datasets that have more heterogeneous electrophysiological properties and artifacts compared to those from adults. We avoided a single-channel approach owing to the inherent properties of pediatric recordings. We even minimized preprocessing steps without the involvement of additional time series filtering and artifact rejection [28]. Our performance, however, was limited to the three-stage classification. Hence, we need to expand our approach to four- or five-stage classifications including N3 and REM stages with additional pediatric EEG datasets. Handling raw time series only in fully end-to-end DNN-based classifiers for automatic sleep staging must be the most convenient strategy. We proposed a slight modification of the fully end-to-end classifier based on the addition of frequency domain information to improve our performance. Instead of complex hand-engineered features, including time-frequency distributions [18], [27], entropy [23], and functional connectivity [20], we selected the IF which was simply obtained from time series, to maintain the nature of the end-to-end classifier as much as possible.

B. INFLUENCES OF THE INPUT LENGTH, NUMBER OF CHANNELS, AND FREQUENCY DOMAIN INFORMATION

The used of a small number of channels is possibly advantageous in terms of convenience for EEG recordings for mobile or wearable devices. Accordingly, DNN-based sleep staging methods with the use of a single-channel have been adopted in many recent studies [21], [24]–[26], [28]. Multi-channel DNN-based sleep staging methods have also been adopted in many recent studies [17], [18], [22], [27], and have mainly used two to six channels depending on the number of available datasets. Chambon *et al.* [22] demonstrated the performance enhancement as a function of the number of channels but saturated beyond a certain point. Given that no recent DNN-based sleep staging studies exist on pediatric EEG datasets, we exploited multi-channel sleep staging approaches with six channels as generally used in PSG studies, and with 19 channels covered all the scalp regions. In the six-channel case, we conjectured that input signals with a length longer than 20 s and frequency domain information are unable to guarantee performance enhancement. Increasing the input length and adding more channels are able to guarantee performance enhancement (30 s and 19 channels were used in our study), unless the two conditions are applied independently, which is partially in line with the observation by Chambon *et al.* [22] for the influence of channel selection. The same author showed that additional modalities including EOG and EMG boosted the performance. However, in this study we solely considered EEG recordings to investigate pure contributions of their temporal and spatial variations to sleep-stage classification. The enhancement of the overall accuracy according to the increase in the input length showed different patterns when six and 19 channels were used, regardless of the types of input signals. This implied that an appropriate combination of temporal and spatial information would be required as an input to the HNN-based classifier for automatic sleep staging.

Many recent studies exploited frequency domain features in sleep staging, including PSD estimates, wavelet entropy, cross-frequency coupling, IF, and time-frequency distributions, with traditional machine learning-based [9], [11]–[15], [20] and DNN-based [18], [23] methods. For our HNN approach, we investigated the use of IF to maintain information based on the temporal variation instead of the PSD estimates, and to avoid computational complexity instead of the entropy, cross-frequency coupling, or time-frequency distributions. In the 19-channel case, the performance was favorably enhanced based on the combination of the time series with IF with an input length of 30 s. This was particularly evident in N1 stage classification. We speculated that the IF in input signals provided considerable frequency domain information for the characterization of each sleep stage [12] for N1 stage classification when the input signals contained sufficient spatial and temporal information, e.g., when 19 channels were used with 30 s signals. The classification accuracy for the W and N2 stages were maintained, or were even slightly decreased based on the addition of IF to the time series in the 19-channel cases with 30 s and 20 s input signals. This suggested that time domain information was sufficient to discern the two types of stages. Additionally, we observed that the 19-channel time series with a 30 s input length in combination with its PSD estimates led to poorer performance owing to the loss of temporal variation in the frequency domain information.

C. POTENTIAL FOR PERFORMANCE IMPROVEMENT VIA DATA ACCUMULATION

The N1 stage appears during the transition from the awake to the sleep states. Therefore, it is difficult and important for a classifier to yield a good performance on N1 stage classification [42]. In addition, given that deep learning is a data driven method, it is expected to yield a better performance for larger training data sizes. We conducted additional experiments to analyze the influence of dataset sizes on the performance in N1 classification of the proposed model compared to shallow models using BP as an input, and especially RF and GB.

Figure 3 shows the F1 scores of class N1 for each tested set according to the ratio, whereby the bold line is the fitted response based on logistic regression. HNN-multi shows

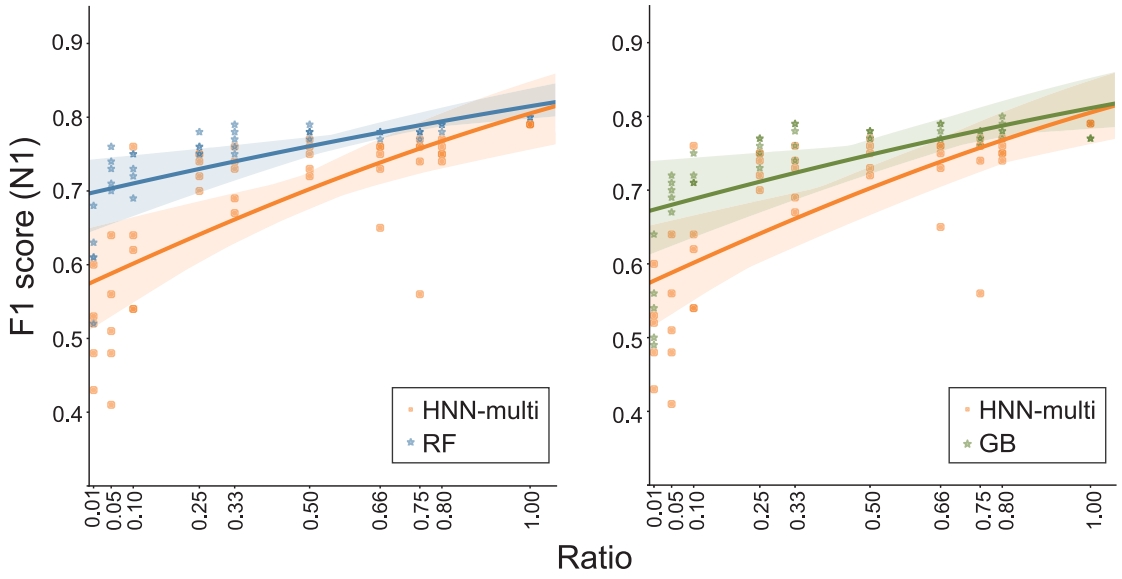


FIGURE 3. F1 score of class N1 according to the dataset size where x-axis is the ratio of subsets to the full dataset. The value of unity in ratio indicates the use of a full training data. To consider the variance in reference to the choice of the subsets, we repeated the same experiment with five different subsets.

better N1 classification performance than GB with full datasets. Furthermore, the proposed model shows continuous increase as a function of the dataset size. By contrast, shallow models performances are saturated in spite of data size increase. We can expect that our model might have a potential to yield better performance in N1 classification when trained with more data via accumulation, which cannot be expected in the cases of conventional machine learning models with feature extraction.

D. CLINICAL IMPLICATIONS

Unlike previous DNN-based sleep staging studies using publicly available PSG datasets, we exploited routine EEG dataset of pediatric patients. Sleep staging is the critical first step for the identification of vigilance and subsequent clinical EEG interpretations. The performance of our HNN-based classifier, which is equivalent to or outperforms previous DNN-based approaches, suggested that our method can be applicable to clinical pediatric automatic sleep staging with PSG datasets and routine EEG recordings, and even in the cases of long-term monitoring. To the best of our knowledge, this is the first DNN-based sleep staging study using pediatric scalp EEG recordings. It has been reported that human EEG signals in pediatric ages undergo considerable variations as the brains of the children mature [43]. In spite of the highly variable pediatric EEG signatures, our performance is sufficient compared with that of recent DNN-based classifiers using adult EEG datasets [17], [21], [22], [25], [26], [28], at least in three-stage classification. Consequently, we can suggest that our HNN-based approach shows clinical feasibility for automatic sleep staging in pediatric ages, even earlier than twelve-year old, i.e., ages lower than the mean ages of this study. The AASM criteria for manual sleep staging rely on specific rules based on certain EEG features and their qualities and quantities. To apply these rules more objectively for clinical sleep staging, we may pursue additional studies to achieve performance enhancements with an input length shorter than the AASM criteria and with the use of additional information, e.g., 10 s time series with 19 channels, 20 s time series with 19 channels, and 20 s time series in combination with IF and 19 channels. The overall accuracies of these three subcases in this study were 91.19%, 91.25%, and 91.11%, respectively, and were nearly equivalent to or outperformed previous DNN-based studies which were based on the AASM criteria with a single-channel [21], [24]–[26], [28], limited PSG channels [17], [18], [22], [27], and whole-brain channels [22].

VI. CONCLUSION

Here, we demonstrated the influences of the length of input signals, number of channels, and types of input signals on the HNN-based automatic sleep-stage classification for pediatric scalp EEG datasets in the time and frequency domains. To boost the performance in three-stage classification, we proposed the use of input signals based on the following conditions: increase of the input length, addition of more channels, and combination of time with temporally variable frequency domain information. We also showed that our HNN-based approach outperformed traditional feature-based machine learning methods, which shed light on the clinical feasibility of DNN-based automatic sleep staging in the field of pediatric neurology. Our study was limited given that sleep stages were limited to three classes and no other electrophysiological recordings, i.e., EOG, EMG and EKG, were considered as input signals. We plan to expand our

sleep stages up to five classes, including N3 and REM stages. Additionally, we can feed additional recording modalities into our HNN-based classifier as input signals to examine supplementary electrophysiological features in each sleep stage.

ACKNOWLEDGMENT

(Yonghoon Jeon and Siwon Kim contributed equally to this work.)

REFERENCES

[1] T. E. Scammell, E. Arrigoni, and J. O. Lipton, "Neural circuitry of wakefulness and sleep," *Neuron*, vol. 93, no. 2, pp. 747–765, 2017.
[2] B. A. Mander, J. R. Winer, and M. P. Walker, "Sleep and human aging," *Neuron*, vol. 94, no. 1, pp. 19–36, 2017.
[3] K. Wulff, S. Gatti, J. G. Wettstein, and R. G. Foster, "Sleep and circadian rhythm disruption in psychiatric and neurodegenerative disease," *Nature Rev. Neurosci.*, vol. 11, no. 2, p. 589, 2010.
[4] R. B. Berry, R. Brooks, C. E. Gamaldo, S. M. Harding, C. L. Marcus, and B. V. Vaughn, "The AASM manual for the scoring of sleep and associated events: Rules, terminology and technical specifications," Amer. Acad. Sleep Med., Darien, IL, USA, version 2.0, 2012.
[5] S. Seifpour, H. Niknazar, M. Mikaeili, and A. M. Nasrabadi, "A new automatic sleep staging system based on statistical behavior of local extrema using single channel EEG signal," *Expert Syst. Appl.*, vol. 104, pp. 277–293, Aug. 2018.
[6] M. Diykh and Y. Li, "Complex networks approach for EEG signal sleep stages classification," *Expert Syst. Appl.*, vol. 63, pp. 241–248, Nov. 2016.
[7] A. R. Hassan, M. Imamul, and M. 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.
[8] P. Chriskos, C. A. Frantzidis, P. T. Gkivogkli, P. D. Bamidis, and C. Kourtidou-Papadeli, "Achieving accurate automatic sleep staging on manually pre-processed EEG data through synchronization feature extraction and graph metrics," *Frontiers Hum. Neurosci.*, vol. 12, p. 110, Mar. 2018.
[9] T. Lajnef, S. Chaibi, P. Ruby, P.-E. Aguera, J.-B. Eichenlaub, M. Samet, A. Kachouri, and K. Jerbi, "Learning machines and sleeping brains: Automatic sleep stage classification using decision-tree multi-class support vector machines," *J. Neurosci. Methods*, vol. 250, pp. 94–105, Nov. 2015.
[10] 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.
[11] P. Memar and F. Faradji, "A novel multi-class EEG-based sleep stage classification system," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 1, pp. 84–95, Jan. 2018.
[12] F. Karimzadeh, R. Boostani, E. Seraj, and R. Sameni, "A distributed classification procedure for automatic sleep stage scoring based on instantaneous electroencephalogram phase and envelope features," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 2, pp. 362–370, Feb. 2018.
[13] V. Bajaj and R. B. Pachori, "Automatic classification of sleep stages based on the time-frequency image of EEG signals," *Comput. Methods Programs Biomed.*, vol. 112, no. 3, pp. 320–328, 2013.
[14] L. Fraiwan, K. Lweesy, N. Khasawneh, H. Wenz, and H. Dickhaus, "Automated sleep stage identification system based on time–frequency analysis of a single EEG channel and random forest classifier," *Comput. Methods Programs Biomed.*, vol. 108, no. 1, pp. 10–19, 2012.
[15] S. I. Dimitriadis, C. Salis, and D. Linden, "A novel, fast and efficient single-sensor automatic sleep-stage classification based on complementary cross-frequency coupling estimates," *Clin. Neurophysiol.*, vol. 129, no. 4, pp. 815–828, 2018.
[16] R. Boostani, F. Karimzadeh, and M. Nami, "A comparative review on sleep stage classification methods in patients and healthy individuals," *Comput. Methods Programs Biomed.*, vol. 140, no. 1, pp. 77–91, 2017.
[17] A. Malafeev, D. Laptev, S. Bauer, X. Omlin, A. Wierzbicka, A. Wichniak, W. Jernajczyk, R. Riener, J. Buhmann, and P. Achermann, "Automatic human sleep stage scoring using deep neural networks," *Frontiers Neurosci.*, vol. 12, p. 781, Nov. 2018.
[18] A. Patanaik, J. L. Ong, J. J. Gooley, S. Ancoli-Israel, and M. W. Chee, "An end-to-end framework for real-time automatic sleep stage classification," *Sleep*, vol. 41, no. 5, p. zsy041, 2018.
[19] Z. Zhang, S. Wei, G. Zhu, F. Liu, Y. Li, X. Dong, C. Liu, and F. Liu, "Efficient sleep classification based on entropy features and a support vector machine classifier," *Physiol. Meas.*, vol. 39, no. 2, 2018, Art. no. 115005.
[20] P. Chriskos, C. A. Frantzidis, P. T. Gkivogkli, P. D. Bamidis, and C. Kourtidou-Papadeli, "Automatic sleep staging employing convolutional neural networks and cortical connectivity images," *IEEE Trans. Neural Netw. Learn. Syst.*, to be published.
[21] A. Sors, S. Bonnet, S. Mirek, L. Vercueil, and J.-F. Payen, "A convolutional neural network for sleep stage scoring from raw single-channel EEG," *Biomed. Signal Process. Control*, vol. 42, pp. 107–114, Apr. 2018.
[22] S. Chambon, M. N. Galtier, P. J. Arnal, G. Wainrib, and A. Gramfort, "A deep learning architecture for temporal sleep stage classification using multivariate and multimodal time series," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 4, pp. 758–769, Apr. 2018.
[23] H. Dong, A. Supratak, W. Pan, C. Wu, P. M. Matthews, and Y. Guo, "Mixed neural network approach for temporal sleep stage classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 2, pp. 324–333, Feb. 2018.
[24] N. Michielli, U. R. Acharya, and F. Molinari, "Cascaded LSTM recurrent neural network for automated sleep stage classification using single-channel EEG signals," *Comput. Biol. Med.*, vol. 106, pp. 71–81, Mar. 2019.
[25] 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.
[26] E. Bresch, U. Großekathöfer, and G. Garcia-Molina, "Recurrent deep neural networks for real-time sleep stage classification from single channel EEG," *Frontiers Comput. Neurosci.*, vol. 12, p. 85, Oct. 2018.
[27] S. Biswal, H. Sun, B. Goparaju, M. B. Westover, J. Sun, and M. T. Bianchi, "Expert-level sleep scoring with deep neural networks," *J. Amer. Med. Inform. Assoc.*, vol. 25, no. 2, pp. 1643–1650, 2018.
[28] S. Mousavi, F. Afghah, and U. R. Acharya, "SleepEEGNet: Automated sleep stage scoring with sequence to sequence deep learning approach," 2019, *arXiv:1903.02108*. [Online]. Available: https://arxiv.org/abs/1903.02108
[29] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, p. 436, 2015.
[30] A. H. Ansari, O. De Wel, M. Lavanga, A. Caicedo, A. Dereymaeker, K. Jansen, J. Vervisch, M. De Vos, G. Naulaers, and S. V. Huffel, "Quiet sleep detection in preterm infants using deep convolutional neural networks," *J. Neural Eng.*, vol. 15, no. 2, 2018, Art. no. 066006.
[31] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal. I. Fundamentals," *Proc. IEEE*, vol. 80, no. 4, pp. 520–538, Apr. 1992.
[32] L. A. Jeni, J. F. Cohn, and F. De La Torre, "Facing imbalanced data–recommendations for the use of performance metrics," in *Proc. Humaine Assoc. Conf. Affect. Comput. Intell. Interact.*, 2013, pp. 245–251.
[33] J. Cohen, "A coefficient of agreement for nominal scales," *Edu. Psychol. Meas.*, vol. 20, no. 1, pp. 37–46, 1960.
[34] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 9, pp. 1263–1284, Sep. 2008.
[35] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in PyTorch," in *Proc. NIPS AutodiffWorkshop*, 2017.
[36] S. J. Reddi, S. Kale, and S. Kumar, "On the convergence of adam and beyond," 2019, *arXiv:1904.09237*. [Online]. Available: https://arxiv.org/abs/1904.09237
[37] P. D. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Trans. Audio Electroacoust.*, vol. 15, no. 2, pp. 70–73, Jun. 1967.
[38] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
[39] L. Mason, J. Baxter, P. L. Bartlett, and M. Frean, "Boosting algorithms as gradient descent," in *Proc. 12th Int. Conf. Neural Inf. Process. Syst. (NIPS)*. Cambridge, MA, USA: MIT Press, 1999, pp. 512–518.
[40] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.

[41] H.-S. Choi, S. Kim, J. E. Oh, J. E. Yoon, J. A. Park, C.-H. Yun, and S. Yoon, "Xgboost-based instantaneous drowsiness detection framework using multitaper spectral information of electroencephalography," in *Proc. ACM Int. Conf. Bioinf., Comput. Biol., Health Inform.*, 2018, pp. 111–121.
[42] J. Virkkala, J. Hasan, A. Värri, S.-L. Himanen, and K. Müller, "Automatic sleep stage classification using two-channel electro-oculography," *J. Neurosci. Methods*, vol. 166, no. 2, pp. 109–115, 2007.
[43] D. L. Schomer and F. H. L. da Silva, *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Philadelphia, PA, USA: Lippincott Williams & Wilkins, 2012.



YONGHOON JEON received the B.E. degree in computer science and the B.A. degree in management as a double-major from Handong Global University. He is currently a Researcher with the Healthcare ICT Research Center, Seoul National University Bundang Hospital, South Korea. His research interests include machine learning and deep learning.



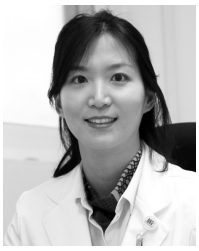
SIWON KIM received the B.S. degree in electrical and computer engineering from Seoul National University, Seoul, South Korea, in 2018, where she is currently pursuing the integrated M.S./Ph.D. degree in electrical and computer engineering. Her research interests include artificial intelligence, deep learning, and biomedical applications.



HYUN-SOO CHOI received the B.S. degree in computer and communication engineering as first major and brain and cognitive sciences as second major from Korea University, South Korea, in 2013. He is currently pursuing the integrated M.S./Ph.D. degree in electrical and computer engineering with Seoul National University, South Korea. His research interests include biomedical signal processing, machine learning, and deep learning.



YOON GI CHUNG received the B.S. degree in electronic engineering and life science as a double-major from Sogang University, South Korea, in 2003, the M.S. degree in biomedical engineering from Drexel University, PA, USA, in 2007, and the Ph.D. degree in brain and cognitive engineering from Korea University, South Korea, in 2014. He is currently a Researcher with the Healthcare ICT Research Center, Seoul National University Bundang Hospital, South Korea. His research interests include biomedical signal processing and statistical machine learning.



SUN AH CHOI received the medical degree and the M.S. degree in medicine from the Seoul National University College of Medicine, Seoul, South Korea, in 2010 and 2017, respectively. She completed her internship and residency in pediatrics at Seoul National University Children's Hospital, Seoul. She is board certified with special qualifications in pediatric neurology, pediatrics, and neurophysiology. She is currently an Assistant Professor in pediatrics with Dankook University Hospital, Cheonan, South Korea. Her research interests include epilepsy, epilepsy surgery, and electroencephalograms.



HUNMIN KIM received the M.D. and Ph.D. degrees from the Seoul National University College of Medicine. He is currently a Pediatric Neurologist with Seoul National University Bundang Hospital. His main research interests include clinical electrophysiology, epilepsy medical treatment, and surgery for intractable epilepsy.



SUNGROH YOON (S'99–M'06–SM'11) received the B.S. degree in electrical engineering from Seoul National University, South Korea, in 1996, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, CA, USA, in 2002 and 2006, respectively. He held research positions at Stanford University and Synopsys, Inc., Mountain View, CA, USA. From 2006 to 2007, he was with Intel Corporation, Santa Clara, CA, USA. He was an Assistant Professor with the School of Electrical Engineering, Korea University, from 2007 to 2012. He is currently an Associate Professor with the Department of Electrical and Computer Engineering, Seoul National University. His research interest includes machine learning and its applications to bioinformatics. He was a recipient of the 2013 IEEE/IEEK Joint Award for Young IT Engineers.



HEE HWANG received the M.D. degree from the Seoul National University College of Medicine. He is a Pediatric Neurologist and a Chief Information Officer with Seoul National University Bundang Hospital (SNUBH).



KI JOONG KIM received the M.D. and Ph.D. degrees from the Seoul National University College of medicine. He is a Pediatric Epileptologist and a Professor with the Department of Pediatrics, Seoul National University Children's Hospital and Seoul National University College of Medicine.