IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

SPECIAL SECTION ON SMART HEALTH SENSING AND COMPUTATIONAL INTELLIGENCE: FROM BIG DATA TO BIG IMPACTS

Received August 27, 2019, accepted September 16, 2019, date of publication October 7, 2019, date of current version October 21, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2946053

Learning-Based Instantaneous Drowsiness Detection Using Wired and Wireless Electroencephalography

HYUN-SOO CHOI[1], SEONWOO MIN[1], SIWON KIM[1], HO BAE[2], JEE-EUN YOON[3], INHA HWANG[3], DANA OH[3], CHANG-HO YUN[3], SUNGROH YOON[1,2,4], (Senior Member, IEEE)

[1]Department of Electrical and Computer Engineering, Seoul National University, Seoul 08826, South Korea
[2]Interdisciplinary Program in Bioinformatics, Seoul National University, Seoul 08826, South Korea
[3]Department of Neurology, Seoul National University Bundang Hospital, Seongnam 13620, Korea
[4]ASRI, INMC, ISRC, and Institute of Engineering Research, Seoul National University, Seoul 08826, South Korea

Corresponding authors: Chang-Ho Yun (ych333@gmail.com) and Sungroh Yoon (sryoon@snu.ac.kr)

This work was supported in part by the Seoul National University Bundang Hospital Research Fund under Grant 13-2016-009, in part by the National Research Foundation of Korea (NRF) through the Ministry of Science and ICT under Grant 2018R1A2B3001628, in part by the Korea Health Technology Research and Development Project under the Korea Health Industry Development Institute (KHIDI) through the Ministry of Health and Welfare under Grant HI15C3224, and in part by the Brain Korea 21 Plus Project in 2019.

ABSTRACT Instantaneous drowsiness (*i.e.,* lapse or micro-sleep) during various activities such as driving or construction causes enormous socioeconomic losses. Thus, a virtuous cycle system that monitors a subject's drowsiness can improve work efficiency and safety. We propose a novel framework to detect instantaneous drowsiness with only a two-second length of electroencephalography (EEG). To achieve reliable performance, we use multitaper power spectral density for feature extraction along with extreme gradient boosting as a machine learning classifier. In addition, we introduce a novel phenotype labeling of instantaneous drowsiness by combining both task dependent and independent measures of alertness (psychomotor vigilance task and electrooculography technique, respectively). The results show that our techniques outperform others used in previous studies. We also identified which spectral components (θ, α, and γ) and channels (Fp1, Fp2, T3, T4, O1, O2, and electrocardiogram) play important roles in our drowsiness detection framework. To verify the applicability for a mobile environment, we implemented our framework on a wireless EEG, as well as on a wired EEG. We hereby present our successful results.

INDEX TERMS Drowsiness, lapse, electroencephalography, wireless electroencephalography.

I. INTRODUCTION

A sufficient amount and good-quality sleep is directly related to cognitive function. Surprisingly, however, over 30% of adults are chronically sleep-deprived, sleeping less than seven hours [1]–[5]. They commonly suffer from major sleep disorders such as insomnia or sleep apnea, which usually results in extreme daytime drowsiness [4], [6], [7]. The extreme daytime drowsiness is associated with lowered attention, which causing considerable socioeconomic burden to the community [8]–[10]. It hinders work productivity [11], lowers academic achievements [12], [13], and increases the risks of traffic or workplace accidents [14], [15]. Therefore, a virtuous cycle system that can monitor drowsiness or attention and provide proper feedback is of vital importance both for improving work efficiency and for the safety of our society.

We can categorize previous drowsiness detection approaches into two types, *i.e.*, task performance and biosignal-based methods. One representative example of the task performance-based method is vehicle motion (VM) monitoring [16]–[19]. However, its usage is restricted to a driving condition only. Another representative method is psychomotor vigilance task (PVT) [20]–[23], where subjects are asked to respond to certain stimuli (visual or audio) as fast as possible. However, the limitation of PVT is that the subject must stop their ongoing task. Electrooculography (EOG) monitoring [19], [24]–[28] is one of the biosignal-based methods. It tracks eye movements or blinking patterns (*e.g.* frequency and speed) to detect drowsiness using a video camera or an infrared device. The limitation of the

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

EOG-based approaches is that they only capture secondary eye responses caused by the homeostatic or circadian sleep drive in the central nervous system. Inevitably they have a low temporal resolution, and are influenced by environmental factors such as wind, temperature, and humidity.

Current algorithms for drowsiness detection are mostly based on electroencephalography (EEG) monitoring. These studies measure electrical activities of the brain associated with drowsiness, extract meaningful features from EEG, and use a classifier to distinguish various states of one's alertness and drowsiness [29]–[31]. Some previous works have adopted machine learning-based classifiers and have shown promising classification results [32]–[43]. However, they mainly focused on classifying long-term states such as conditions before or after driving, and they fail to capture more *instantaneous drowsiness* (*i.e.* lapse) which may cause severe accidents.

In this work, to address the limitations of previous works, we propose a novel framework to detect *instantaneous drowsiness* using a short time segment ($\sim$ 2 seconds) of EEG. The main contributions of this paper, as an extension of our previous work [44], are as follows:

- We define a novel phenotype labeling method for *instantaneous drowsiness* by combining the advantages of PVT (as a standard reference) and EOG monitoring (as a task-independent measurement).
- We propose a novel framework outperforming the previous approaches by adopting multitaper power spectral density (MPSD) for EEG feature extraction and extreme gradient boosting (XGBoost) as a classifier.
- We identify key frequency bands and channels for drowsiness detection.
- We demonstrate that using only seven channels (Fp1, Fp1, T3, T4, O1, O2, and ECG) can provide comparable performance to using the original twenty with less than 2% accuracy degradation.
- We verify the applicability of the proposed framework for a mobile environment by using a wireless EEG with dry-sensors as well as a wired EEG with wet-sensors.

II. RELATED WORKS

A. TASK PERFORMANCE-BASED DROWSINESS DETECTION METHODS

The VM monitoring module [16]–[19], and the PVT [20]–[22] determine the drowsiness level by measuring task performance. The VM monitoring modules receives input from driving directions and from changes in lane-keeping. It is vulnerable to weather, road condition, and vehicle type. The PVT is a well-established measure of alertness or sustained attention. For the standard ten-minutes PVT, the subjects are required to click a button with the dominant hand's thumb as soon as possible when the visual signals are presented in random intervals (2–10 seconds) [20], [23]. Response time is a validated indicator of the alertness level. The PVT can measure changes in alertness caused by sleep disorders and deprivation [20]–[23]. Sleep deprivation leads to a fluctuation in sustained attention because of the interaction of involuntary sleep-initiating and counteracting wake-maintaining systems, thus resulting in lapses (errors of omission to respond for given stimuli). However, to complete the given task, subjects need to interrupt the ongoing task, thus hindering the work-continuity. In this study, the PC-PVT platform (Biotechnology High Performance Computing Software Applications Institute, http://bhsai.org/software/pcpvt, MD, USA) is adopted [45] (Section III-C). Both PC-PVT and standard PVT have the same functionality. PC-PVT uses a personal computer, whereas PVT uses a specific hardware. The PC-PVT data are used to label drowsiness for EEG segments. The labeled data are then utilized to train the supervised classification-based drowsiness detection model.

B. EOG-BASED DROWSINESS DETECTION METHODS

EOG-based drowsiness detection uses a video camera to analyze eye movement markers such as speed, frequency, blinking, and winding. Daytime drowsiness is closely associated with several ocular parameters (*e.g.,* slow movements, increased closure time, and increased blinking frequency) [19], [24], [26]–[28]. R100 (Phasya, Belgium) is a recently developed EOG-based method [46]. The R100 exploits glasses equipped with a high-speed camera to sense eye and eyelid movements. It can continuously monitor the level of drowsiness with minimal or no disruption of the ongoing task, and provide a task-independent measure of drowsiness. In this study, we use the R100 device to acquire a long length of drowsiness information. The obtained information is used to define the drowsiness state label discussed in Section III-C.

C. EEG-BASED DROWSINESS DETECTION METHODS

EEG changes are closely related to alertness fluctuations. Regarding drowsiness, Putilov and Donskaya [47] reported that when the subjects's eyes are open, the power of the alpha and theta bands increase, and with eye closure only the theta band power increases. Alloway *et al.* [48] reported that the alpha band power increases when the eyes are open but decreases when they close.

In the previous studies, the drowsiness label was variously defined. The methods for labeling can be categorized into three types. The first category is a questionnaire-based method such as the Epworth [49] and Karolinska sleepiness scale [47]. A simple questionnaire such as ''how sleepy are you now?'' is given to the subject. The output is hardly objective. The second approach is task performance evaluation that uses VM detection modules [32], [33], [38], [50], [51] or response time to specific stimuli [37], [51], [52]. The third type is EOG-based method. Li and Chung [53] defined the label with EOG information but only used the parameter of eyelid closure ratio.

To the best of our knowledge, our drowsiness labeling method is the first approach that utilizes both R100 and PVT. R100 provides a task-independent and long-term

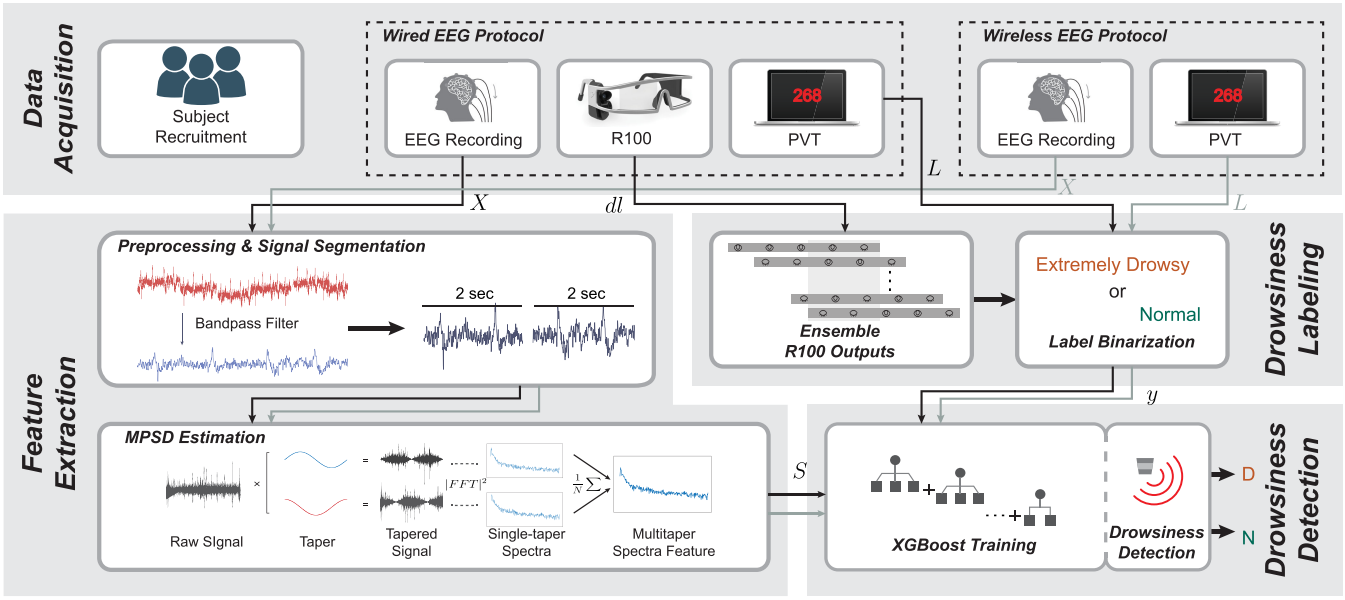


FIGURE 1. The Framework is composed of four main steps: Data acquisition, feature extraction, drowsiness labeling, and drowsiness detection. For wireless EEG protocol, the drowsiness level was evaluated only with PVT (please see Section III-E for the details).

measurement of drowsiness, and PVT sets the references for extreme sleepiness.

D. FEATURE EXTRACTION FOR EEG SIGNALS

Many types of feature extraction techniques have been proposed to retrieve useful EEG information and apply it to drowsiness detection studies. The most widely used feature is band power (BP). Especially theta ($4-8$ Hz) and alpha ($8-16$ Hz) bands have played a key role [37], [38], [54], [55]. To achieve more detailed frequency information than BP, some studies [41], [50], [52] have used power spectral density (PSD) estimation or wavelet decomposition. However, the PSD-based approaches suffer from high bias and variance in estimation. A single-taper PSD (SPSD) method was developed to address high bias issues [56]. Nonetheless, the SPSD technique was not able to fix the high variance in estimation. To solve this issue, a multitaper PSD (MPSD) method has been developed [57]. MPSD allows precise estimation by aggregating multiple independent SPSDs, thus outperforming SPSD for estimating the sleep stages [58]. To the best of our knowledge, this is the first study that applies MPSD for detecting drowsiness with EEG.

E. MACHINE LEARNING METHODS FOR DROWSINESS DETECTION

Many studies based on machine learning techniques have been proposed for detecting drowsiness. Linear regression was first applied for drowsiness detection [37]. Since then, several studies have used artificial neural networks (ANNs) [34], [41], [51], [52] and support vector machines (SVMs) [33], [51], [54]. In our study, we use the extreme gradient boosting (XGBoost) method [59], which offers the following benefits: ease of use, scalability, accuracy, and computational efficiency. In addition, it has good records on recent machine learning competitions [60]. Furthermore, we can estimate the importance of each input feature by using the branch gain in XGBoost. This information can suggest which EEG frequency bands and channels are relevant to the drowsiness level, and guide the target of the electrode application and the frequency band analysis, thereby enhancing performance and reducing resource inputs in future studies.

III. METHODS

As depicted in Figure 1, the proposed framework includes four steps: data acquisition, feature extraction, drowsiness labeling, and drowsiness detection. For the first step, data acquisition, we acquire the EEG signals from the recruited subjects. Furthermore, we perform an R100 and a PC-PVT (see section III-A). As the second step, we extract the MPSD [57] feature from the preprocessed EEG segments (see section III-B). In the next step, the instantaneous drowsiness label is defined for each EEG segment using the R100 ensemble outputs, and is validated by comparing lapse information from PC-PVT (see section III-C). For the last step, XGBoost [59] is trained with the features and labels which are respectively acquired from previous steps. The trained XGBoost decides whether a given EEG segment is acquired during a drowsiness condition or not (see Section III-D).

A. DATA ACQUISITIONS

1) SUBJECTS

We recruited eight healthy subjects (4 men and 4 women; age, 26.8 ± 3.4 years old; body mass index, 20.9 ± 2.1 kg/m^2) who were non smokers and did not have any neurologic diseases and sleep disorders such as obstructive sleep apnea, insomnia, primary hypersomnia, or restless legs syndrome. None took

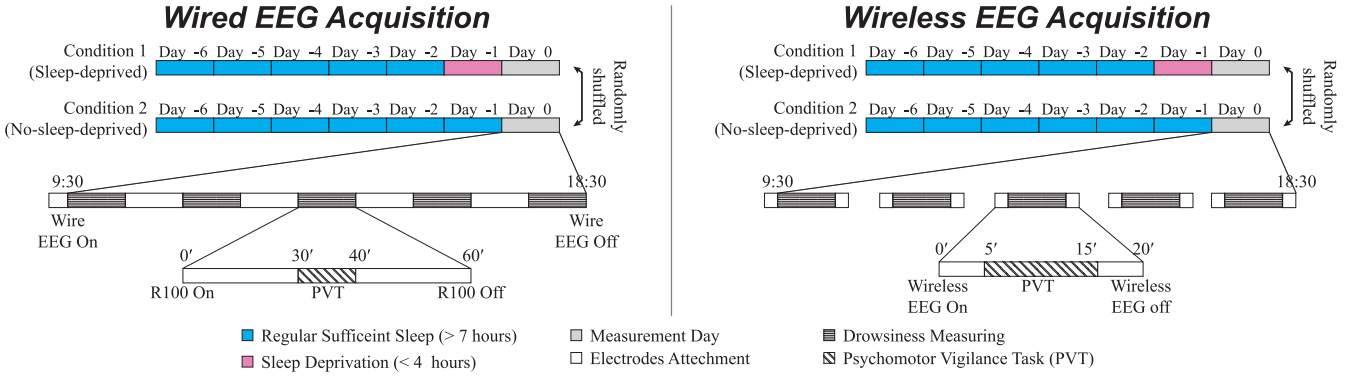


FIGURE 2. Illustration of two EEG acquisition protocol (wired and wireless). There are two random ordered conditions (sleep-deprived and no-sleep deprived) for each protocol. While the wired EEG device is worn for an entire, the wireless EEG device was only worn during the PVT task.

TABLE 1. Sleep time (minutes) for each condition.

No-sleep-deprived		Sleep-deprived	
Night -6 – -2	Night -1	Night -6 – -2	Night -1
441.25±38.50	431.25±31.82*	440.52±44.43	216.25±30.21*

*Sleep time of the night before measurement on each condition showed significant difference ($p < 0.05$).

any medication affecting sleep or alertness, and they reported no involvement in shift work for the previous one-year period and no travel to an area with a different time zone during the month before enrollment. All participants had intermediate chronotype, and slept regularly, sleeping between six and eight hours per day.

The institutional review committee of Seoul national university Bundang hospital and the Seoul national university approved the wired and wireless EEG protocols, respectively. All participants submitted written informed consent.

2) EXPERIMENTAL PROTOCOL

A schematic presentation is shown in Figure 2. We conducted two protocols (wired EEG and wireless EEG), and each protocol consisted of two conditions (sleep-deprived and no-sleep-deprived). Two conditions were randomly ordered, and the inter-condition interval was two weeks. During the study period, subjects were instructed to have regular sleep for more than seven hours per night; however, in the sleep-deprived condition, they were instructed to sleep less than four hours on the night before the EEG acquisition. Each subject's compliance to the given sleep-wake schedule was monitored through daily sleep logs. Sleep duration was successfully controlled as shown in Table 1, 216.3±30.2 minutes on a sleep-restricted day. All the subjects abstained from alcohol and caffeine for at least the 48 hours before and during the day of the EEG measurement.

All EEG measurements were performed from 9:30 AM to 6:30 PM in an isolated space. Subjects were allowed to have regular meals and water, but abstained from drinking caffeine and alcohol, smoking, and napping. The first meal was before 9:30 AM and the lunch was between 12:30 PM and 1:30 PM. The subjects' behaviors were continuously monitored by real-time video.

3) EEG AND DROWSINESS MEASUREMENT

A standard wet-electrode EEG (Beehive Horizon, Grass Technologies, Natus, USA) and a cap-type dry-electrode EEG (Ybrain Inc., Republic of Korea) device were adopted for the wired and wireless EEG protocols, respectively. A total of 19 EEG electrodes were placed according to the standard 10-20 system, and an additional single ECG channel (modified type II lead) was recorded. The recording list is as below:

$$\text{set } \mathcal{C} = \{\text{Fp1, Fp2, F7, F3, Fz, F4, F8, T3, T4,}\\ \text{T5, T6, C3, Cz, C4, P3, Pz, P4, O1, O2, ECG}\}. \quad (1)$$

A referential montage was adopted in which all the electrodes were referenced to the left mastoid electrode (A1).

For the wired EEG protocol, the EEG data were continuously recorded from 9:30 AM to 6:30 PM, and five drowsiness measurements were taken at two-hour interval (Figure 2). Each measurement consisted of two recordings with continuous background EEG monitoring: a 60-minute R100 recording and a ten-minute PVT in the middle of the former. For the wireless protocol, subjects wore the EEG cap every two hours for 20 minutes (Figure 2). As wireless EEG cap interfered the stable application of R100 and the EEG artifacts increased due to the contacts between the temporal EEG electrodes and the temple bows of the R100 eye glasses, only PVT information was used to phenotype drowsiness. For both protocols, all the recording programs were installed on a single laptop to synchronize the time axis for EEG, R100, and PVT data.

B. FEATURE EXTRACTION

1) EEG PREPROCESSING

The jth channel EEG signal x^j is computed by the subtraction from the electric potential i^j to the i^{A1} signal from the left mastoid (channel A1) as below:

$$x^j = i^j - i^{A1}. \quad (2)$$

Other montage (average referential and longitudinal bipolar montage) did not show significant performance difference with the left mastoid reference montage. A 1 Hz low-pass filter and a 50 Hz high-pass filter is set. The sampling rate is 200 Hz.

To prevent drowsiness-related accidents, the length of the EEG signal required for detecting instantaneous drowsiness should be short. In our approach, we split the EEG signal without overlap into a specific length (1–16 seconds) segment. When the size of the window was set to w, the EEG segment could be expressed as follows:

$$x_i^j = [x_{(i-1)*w+1}^j, x_{(i-1)*w+2}^j, \ldots x_{i*w}^j]. \tag{3}$$

2) MULTITAPER SPECTRAL FEATURE EXTRACTION

In this study, multitaper power spectral density (MPSD) [57] based feature extraction is performed on each EEG segment. For the x_i^j in (3), single power spectral density (SPSD) $s_i^j(f)$ [56] at the frequency f is calculated as follows:

$$s_i^j(f) = \Delta t \left| \sum_{k=1}^{w} w_k^{(l)} x_{(i-1)*w+k}^j e^{2\pi k f \Delta t} \right|^2, \tag{4}$$

where w is a taper function and Δt indicates the sampling duration. MPSD is computed by averaging L number of SPSDs generated with orthogonal tapers as follows:

$$\bar{s}_i^j(f) = \frac{1}{L} \sum_{l=1}^{L} s_i^{j(l)}(f). \tag{5}$$

The MPSD of the x_i^j for m frequencies is given as follows:

$$s_i^j = [\bar{s}_i^j(f_1), \bar{s}_i^j(f_2), \ldots, \bar{s}_i^j(f_m)]. \tag{6}$$

The spectral information of n number of channels for the EEG segment X_i is defined as $S_i = [s_i^1, s_i^2, \ldots, s_i^n] \in \mathrm{R}^{m*n}$.

C. DROWSINESS LABELING

Since our approach is based on supervised learning, it is necessary to define a label indicating drowsiness for given EEG segments. To focus on detecting instantaneous drowsiness, we define a label of whether the subject showed lapse or not during PVT measurement. However, EEG with PVT information (50 minutes per day) is not enough to train the model. Thus, to define the label y_i, we aggregate the R100 information (five hours per day) in section III-C1. Then, we find a proper threshold to binary y_i and validate how well y_i is related to the occurrence of the lapse in section III-C2.

1) R100-BASED DROWSINESS LABELING

The EEG segment (1–14 seconds) is relatively short compared to the time required by the R100 (60 seconds); thus, the multiple outputs dl^n of the R100 correspond to a single EEG segment. Therefore, in our work, to increase robustness, we aggregate the R100 outputs into one representative value, as shown in Figure 3. We adopt an averaging method to

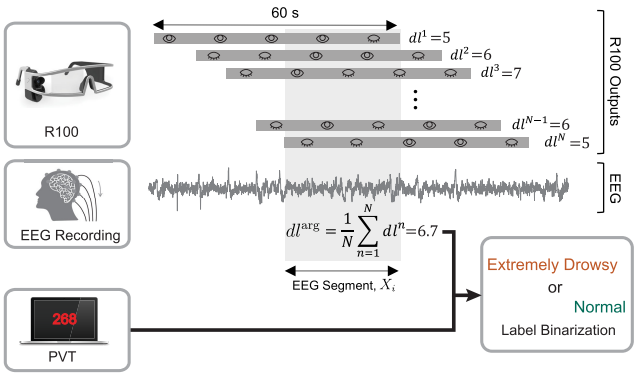


FIGURE 3. Since multiple R100 outputs correspond to given a EEG segment X_i, one representative drowsiness level $dl^{\arg}$ is aggregated by averaging.

define the label $dl_i^{\arg}$ corresponding to the ith EEG segment as follows:

$$dl_i^{\arg} = \frac{1}{N} \sum_{n=1}^{N} dl_i^n. \tag{7}$$

To validate the averaging method, the majority of dls and the first dl^1 output are compared empirically in the experiment.

To binalize the $dl_i^{\arg}$, the following formula is defined:

$$y_i = \begin{cases} 1 & \text{if } dl_i^{\arg} \geq \text{threshold} \\ 0 & \text{if } dl_i^{\arg} < \text{threshold}. \end{cases} \tag{8}$$

$y_i = 1$ represents the instantaneous state and $y_i = 0$ represents the normal state. The following section discusses how to decide the threshold using PVT information.

2) THRESHOLD SEARCHING FOR LABEL BINARIZATION

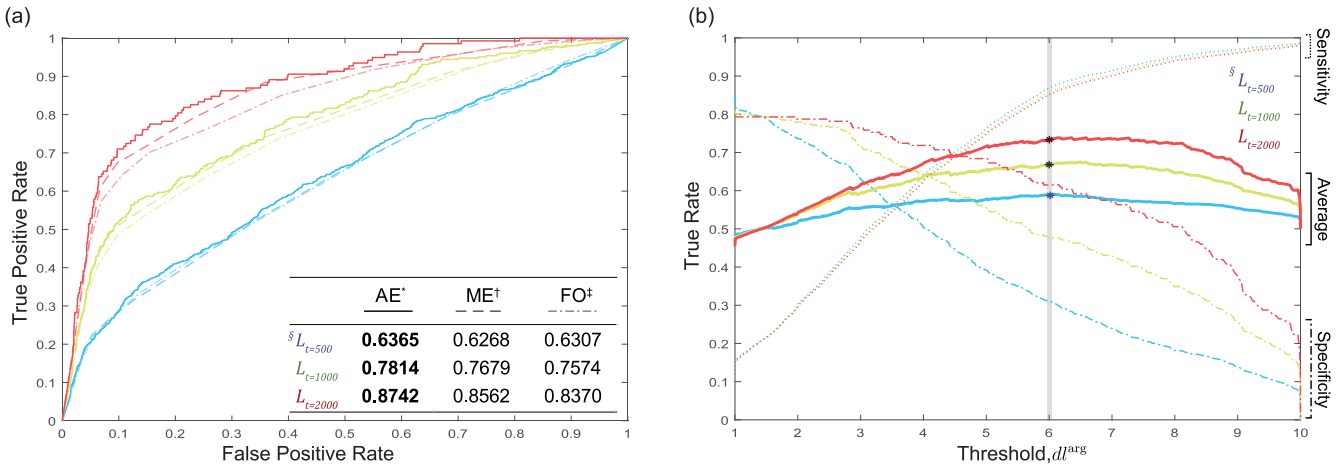
A lapse is a temporary episode of drowsiness which lasts for a second where a subject fails to respond in a certain time (t) [61]. Let RT_i be the PVT response time in X_i, then occurrence of the lapse L_i can be defined as

$$L_i = \begin{cases} 1 & \text{if } \mathrm{RT}_i \geq t \\ 0 & \text{if } \mathrm{RT}_i < t \end{cases}. \tag{9}$$

To determine a precise threshold of (8), we use the method proposed by Francois' [46]. The threshold in (8) is chosen which to minimize the distance between y_i in (8) and L_i in (9). By assuming L_i as target and y_i as a predicted target, we calculate the sensitivity and specificity based on their difference. Thus, the receiver operating characteristic (ROC) curve can be plot by modifying the threshold value. The optimal threshold is determined when the average of the sensitivity and specificity is maximized. Once the threshold is determined, we are able to determine y_i over the entire duration when the subject wears R100.

D. DROWSINESS DETECTION

A binary classifier using the XGBoost is trained for each subject with the feature vector S_i and the corresponding label



*Average Ensemble (Proposed), †Majority Ensemble, ‡First Output, § L in (9) for certain t

FIGURE 4. Evaluation results on three methods of drowsiness level definition. (a) ROC curves and AUC values for the three methods on three values of t. (b) Sensitivity and specificity trends as function of a sweeping AE drowsiness level on three values of t.

y_i respectively as the input and the output. XGBoost is an ensemble method that aggregates a outputs from K number of classification and regression trees (CART) as follows:

$$\hat{y}_i = \sum_{t}^{T} f_t(S_i), \tag{10}$$

where f_k represents the kth tree model. The objective function is given by

$$Obj = \sum_{i}^{M} l(y_i, \hat{y}_i) + \sum_{k}^{K} \Omega(f_k), \tag{11}$$

where $l(\cdot)$ and $\Omega(\cdot)$ are the loss function and the complexity of trees respectively with M training samples. $l(\cdot)$ is the cross-entropy, which is defined as

$$l(y_i, \hat{y}_i) = -y_i \log \hat{y}_i - (1 - y_i) \log(1 - \hat{y}_i). \tag{12}$$

In our problem, the ratio of the negative class (normal state) samples against the positive class (drowsiness) samples is approximately 8 : 2. Since these imbalanced data can cause biased classification [62], by multiplying a weight, we define a cost-sensitive loss $l(y_i, \hat{y}_i)$ as (13)

$$l_c(y_i, \hat{y}_i) = -w_c y_i \log \hat{y}_i - (1 - y_i) \log(1 - \hat{y}_i), \tag{13}$$

where $w_c = (\#\text{ of positive})/(\#\text{ of negative})$.

XGBoost is based on an additive training technique that progressively adds trees to increase the precision of the prediction as

$$\hat{y}_{(t)} = \hat{y}_{(t-1)} + f_t(S), \tag{14}$$

where $t = 1, \ldots, T$, and $\hat{y}_0 = 0$. Since XGBoost is vulnerable to overfitting, the area under curve (AUC) values of the validation sets are used as a criterion for early stopping. The training stops when there is no further improvement after the addion of more than 300 trees. In our work, we used five-fold cross-validation to validate our framework.

By letting j as an index of a single leaf among the K number of leaves in a tree, the object function (11) converges as

$$Obj^* = -\frac{1}{2} \sum_{k=1}^{K} \frac{G_k^2}{H_k + \lambda} + \gamma K, \tag{15}$$

where λ and γ are parameters that control the trade-off relationship between the complexity and accuracy. G and H is defined as

$$G_k = \sum_{i \in I_k} \partial_{\hat{y}_i} l(y_i, \hat{y}_i), \quad H_k = \sum_{i \in I_k} \partial^2_{\hat{y}_i} l(y_i, \hat{y}_i), \tag{16}$$

where I_k denotes a set of samples reaching the kth leaf in each tree. The Obj^* in (15) value measures how well a tree is structured. The optimization of the tree is conducted by expanding the leaves based on information gain (*Gain*) (17) which is defined by

$$\text{Gain} = \frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{G_L^2 + G_R^2}{H_L + H_R + \lambda} - \gamma. \tag{17}$$

Each term respectively denotes the left children score, and the right children score, as well as the score without split. By considering these three cases, a tree is structured to maximize the total gain value. Following a pruning technique, if the gain of each leaf is smaller than γ, tree addition is terminated.

Since each branch of the tree corresponds to one of the feature, we can compute the total gain for a specific feature by summing the gain of each branch. The sum represents the importance of the corresponding feature. In our study, since spectral information of each node is used as a feature, we are able to easily compute the importance of each frequency and each channel.

E. APPLICABILITY IN A WIRELESS EEG ENVIRONMENT

Wireless EEG was additionally acquired to confirm that the framework is applicable in a wireless EEG environment. However, since drowsiness information from the

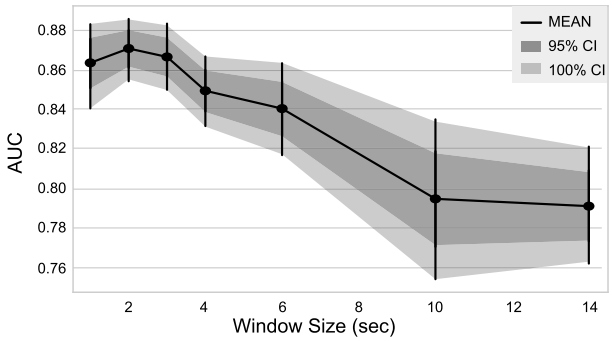


FIGURE 5. The AUC distributions according to window size, w in (3).

R100 was not acquired in the case of wireless EEG, as below, the extreme drowsiness for wireless protocol, y_i^{wl}, was defined directly with the lapse during the PVT tasks.

$$y_i^{\text{wl}} = \begin{cases} 1 & \text{if } \text{RT}_i \geq t \\ 0 & \text{if } \text{RT}_i < t. \end{cases} \tag{18}$$

t is adaptively considered for each subject as $t = \mu + 2 * \sigma$ where μ and σ are respectively the mean and the standard deviation of RT. Except for drowsiness labeling, the rest of the process is the same as with the wired EEG.

IV. RESULTS

A. EVALUATION OF DROWSINESS LABEL

ROC curves of the drowsiness label defined by the three methods (average ensemble, majority ensemble, and first output) in Section III-C are shown in Figure 4 (a). The proposed averaging method shows a high AUC value regardless of t. Therefore, we can conclude that the labeling defined by the averaging aggregation method of R100 is most highly related to lapse during PVT.

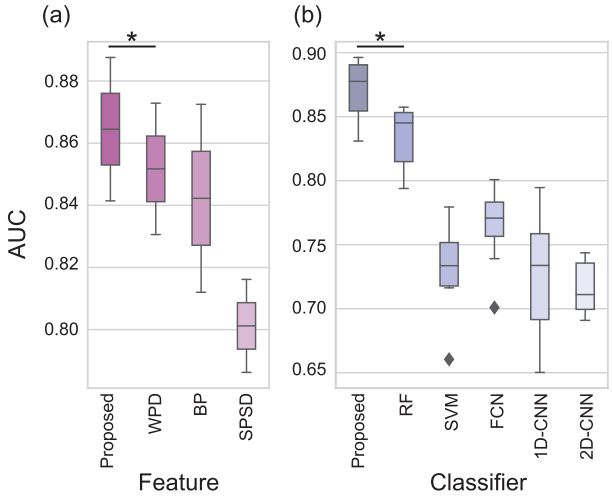
As discussed in Section III-C2, the proper threshold value should be determined. Figure 4 (b) shows the sensitivity and specificity values between L_i in (9) and y_i in (7) for every t. The mean values of sensitivity and specificity are depicted as bold lines. The mean values were maximized on $dl_{ens} = 6$. In other words, if $dl_{ens} \geq 6$ (=threshold), we assumed that the subjects were suffering from instantaneous drowsiness.

B. COMPATIBILITY AS INSTANTANEOUS DROWSINESS DETECTION

To instantaneously detect the drowsiness, the length of the EEG signal should be short. We analyzed the AUC values on various window sizes, w in (3). The highest accuracy was reached with the two-second window size, which is sufficiently short to detect instantaneously. Longer EEG segments resulted in lower accuracy. We think that the chance of a mixed condition of normal and drowsiness states in the long EEG segment might produce a negative effect.

C. COMPARATIVE ANALYSIS

Comparisons between the techniques used in proposed frameworks and other conventional techniques (three methods of feature extraction and five classifiers) were conducted. Figure 6 depicts the distribution of AUC values. By testing Delong's method [63], the statistical significance among the methods was analyzed.



*: AUC showed significant difference ($p < 0.05$) by Delong's method.

FIGURE 6. AUC distributions on (a) features and (b) classifiers.

1) COMPARISON AGAINST OTHER FEATURE EXTRACTION METHODS

Three widely used feature extraction methods were implemented and compared against MPSD. The first method was band power (BP) [54], which uses the power values on specific bands, such as the delta ($0.5 - 4$ Hz), theta ($4 - 8$ Hz), alpha ($8 - 13$ Hz), beta ($13 - 30$ Hz), and gamma ($30 - 50$ Hz) bands. The second was Welch's method based on SPSD [56]. The third method was the wavelet packet decomposition (WPD) [64]. WPD is known to overcome the previous methods' limitation, *i.e.,* the lack of precision with nonstationary signals. For every feature extraction technique, the XGBoost were used. As shown in Figure 6 (a), the method that we used in our work, MPSD, shows the best performance, whereas SPSD shows the worst. The multi-taper technique significantly boosts the performance of extracting the spectral information accurately.

2) COMPARISON AGAINST OTHER CLASSIFIERS

Five widely used machine learning classifiers were implemented and compared. The first classifier was random forest (RF) [65], which is a bagging-based ensemble technique. The second classifier was SVM [66]. The remaining types of classifiers were deep-learning (DL) models [67]. DL models could be implemented into various architectures. We implemented three different architectures: fully connected networks (FCNs), 1D convolutional networks (1D-CNNs), and 2D convolutional networks (2D-CNNs). Due to class imbalance, the training of the SVM and the DL models showed difficulties on convergence. For this reason, we balanced the training data by generating synthetic data using the

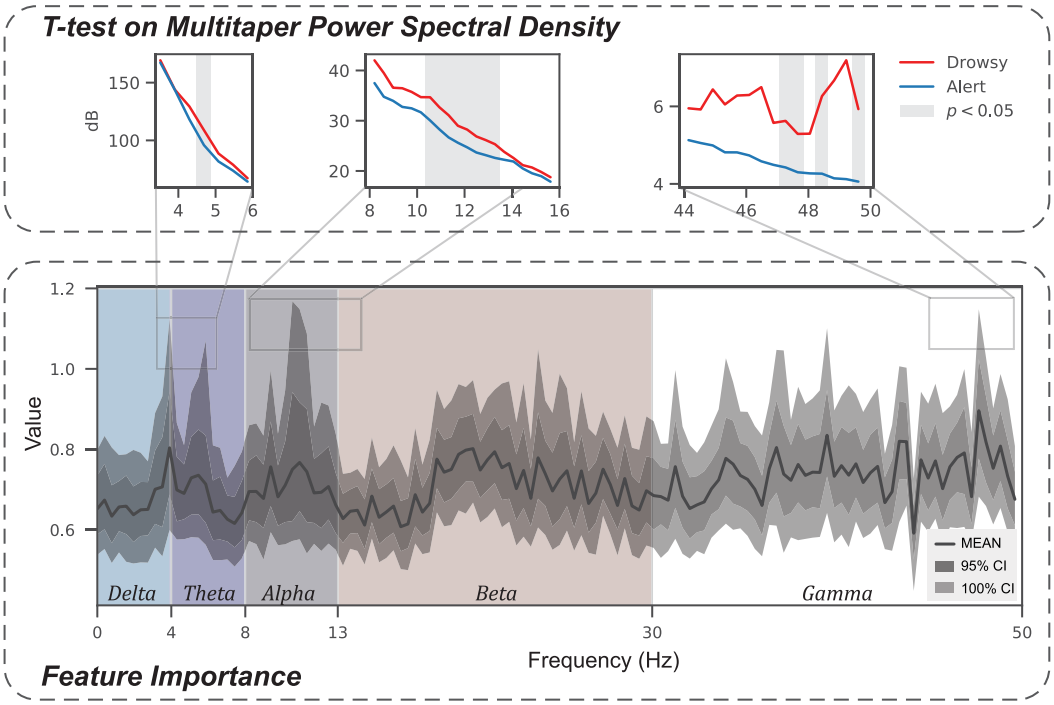


FIGURE 7. Feature importance as a function of frequency. On certain frequencies that show peaks on feature importance, t-test showed significant difference.

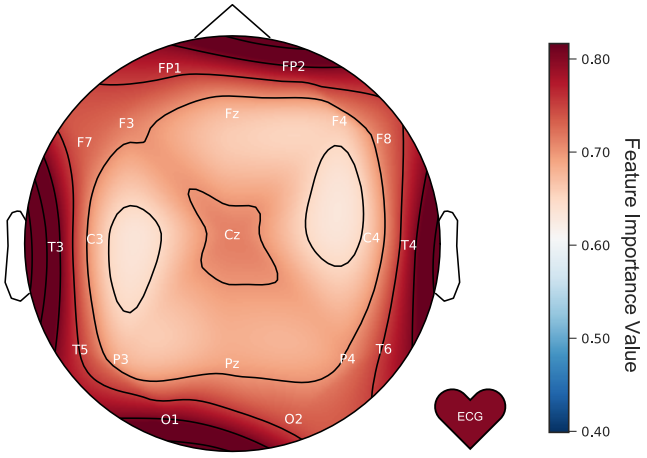


FIGURE 8. Topographic mapping of feature importance value.

synthetic minority over-sampling technique (SMOTE) [68] for the drowsiness class. All the hyper-parameter sets were selected using a greedy search.

As shown in Figure 6 (b), XGBoost which we used in our work, shows the best results. RF-based detection shows relatively good performance against the others. Even though SMOTE was additionally applied as prepossessing to overcome imbalanced data issue, SVM shows relatively low performance with a biased classification on a major class (normal state). Even though we applied various techniques (*e.g.,* such as skip connection, drop-out, and batch normalization), DL models also do not show good performance.

D. FEATURE IMPORTANCE

As described in Section III-D, the importance of the input features was computed using the gain value. The importance values of a specific frequency are shown in Figure 7. Peaks are shown in the theta and the alpha bands, which means that these bands play a key role in decisions. This finding shows consensus with the existing studies (e.g., the alpha attenuation test [48] and the Karolinska drowsiness test [47]). Furthermore, we found that the range $(45 - 50$ Hz) in the gamma band is also important. Figure 8 shows a topological mapping of importance. Seven channels (Fp1, Fp2, T3, T4, O1, O2, and ECG) are the key channels among 10-20 system channels.

E. CHANNEL REDUCTION

If the framework is based on fewer channels, a more lightweight and cost-effective system can be implemented. Detection performance according to various channel combinations is shown as Figure 9. When all channels are used, the best performance is obtained. When using seven selected channels (Fp1, Fp2, T3, T4, O1, O2, and ECG) based on the importance, a small performance drop of -1.80% is observed. As the number of channels decreased, the performance of the system continued to degrade. The performance increased when the ECG channels were added for all the cases. Furthermore, using only the channel of a specific region, the performance was in the order of the ECG-temporal-frontal-occipital region. There was no significant difference in performance between the left hemisphere (Fp1,

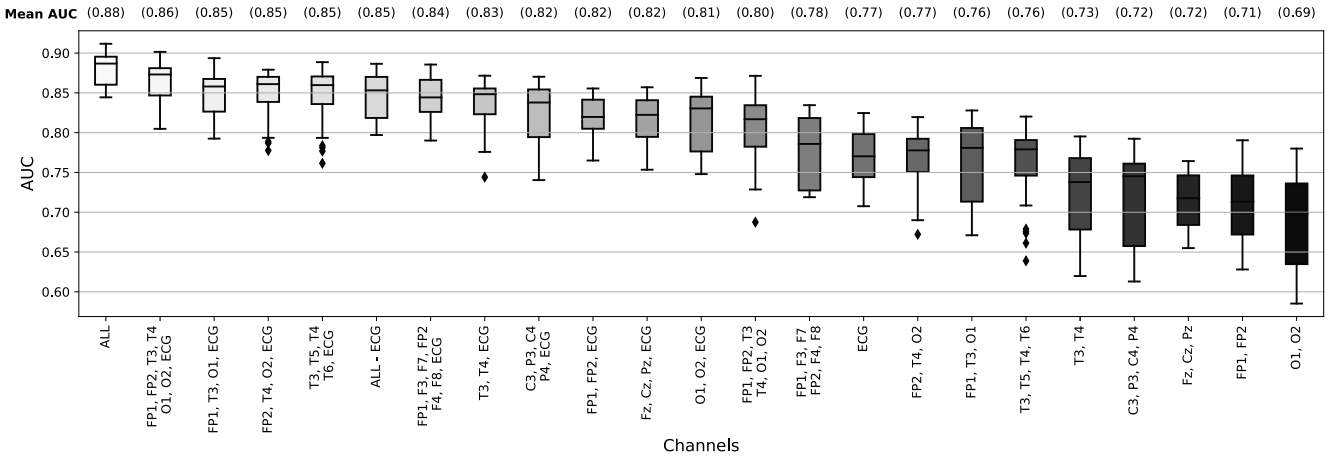


FIGURE 9. Performance degradation according to channel reduction.

TABLE 2. Subject-specific performance of wire EEG and wireless EEG.

	Wire EEG + R100 + MPSD + XGBoost				Wireless EEG + PVT + MPSD + XGBoost			
Subject	Accuracy	AUROC	Sensitivity	Specificity	Accuracy	AUROC	Sensitivity	Specificity
1	0.8033 ± 0.0056	0.8913 ± 0.0038	0.8030 ± 0.0053	0.8034 ± 0.0057	0.8474 ± 0.0374	0.9154 ± 0.0259	0.8133 ± 0.0183	0.8486 ± 0.0383
2	0.7665 ± 0.0139	0.8489 ± 0.0042	0.7678 ± 0.0167	0.7664 ± 0.0137	0.6425 ± 0.0709	0.6766 ± 0.0582	0.6067 ± 0.0683	0.6436 ± 0.0711
3	0.7661 ± 0.0091	0.8461 ± 0.0066	0.7660 ± 0.0087	0.7661 ± 0.0092	0.8200 ± 0.1459	0.8054 ± 0.1941	0.6000 ± 0.4183	0.8247 ± 0.1487
4	0.7980 ± 0.0138	0.8889 ± 0.012	0.7984 ± 0.0141	0.7980 ± 0.0138	0.9132 ± 0.0969	0.9555 ± 0.0401	0.7333 ± 0.4346	0.9167 ± 0.0998
5	0.7428 ± 0.0146	0.8211 ± 0.0147	0.7417 ± 0.0141	0.7429 ± 0.0146	0.6536 ± 0.1215	0.6760 ± 0.1604	0.6000 ± 0.1369	0.6548 ± 0.1216
6	0.8269 ± 0.0080	0.9056 ± 0.0062	0.8267 ± 0.0085	0.8269 ± 0.0079	0.7730 ± 0.0513	0.8238 ± 0.0564	0.7179 ± 0.0542	0.7754 ± 0.0534
7	0.7786 ± 0.0115	0.8609 ± 0.0077	0.7788 ± 0.0119	0.7785 ± 0.0115	0.7706 ± 0.0736	0.8158 ± 0.0756	0.7000 ± 0.0935	0.7727 ± 0.0769
8	0.7986 ± 0.0149	0.8813 ± 0.0106	0.7985 ± 0.0152	0.7986 ± 0.0149	0.7574 ± 0.0905	0.8170 ± 0.0835	0.6933 ± 0.1738	0.7594 ± 0.0940
Average	0.7851 ± 0.0266	0.8680 ± 0.0284	0.7851 ± 0.0267	0.7851 ± 0.0266	0.7722 ± 0.0917	0.8107 ± 0.0988	0.6831 ± 0.0763	0.7745 ± 0.0925

T3, O1, and ECG) and the right hemisphere (Fp2, T4, O2, and ECG).

F. RESULTS ON WIRED AND WIRELESS EEG

As mentioned in Section III-E, in addition to the wired EEG device, we have acquired additional data with the wireless EEG system, which uses dry sensors. The performance of each subject using the wired EEG and wireless EEG is described in Table 2. The performance for each subject represents the mean and standard deviation of the five-fold cross-validation. For the wired EEG, most of the metrics show values above 0.78, which means that the proposed framework can be applied to the actual situation. Additionally, all the metrics have small standard deviations (less than 0.05 among folds and subjects). For the wireless EEG, the AUC values for all subjects showed an average performance degradation of −6.55%. In addition, the standard deviation of performance was higher in the wireless EEG than with the wired EEG. In other words, it can be interpreted that detection is less stable with the wireless EEG than with the wired EEG. This performance degradation could be caused by the instability of the dry sensor or the lack of EEG length used during learning. However, since all indicators show a value of over 0.70, we expect the wireless EEG to be useful for drowsiness detection in mobile a environment.

V. DISCUSSION

We have proposed a novel framework for detecting instantaneous drowsiness, which can be applied regardless of the subject's circumstance. In other words, our framework is able to detect drowsiness in any situation. Once the model is trained, it does not require any extra tasks, such as PVT, to evaluate the drowsiness. Moreover, since our framework uses a short EEG segment of two seconds, it is able to quickly detect and notify instantaneous drowsiness states such as a lapse.

The proposed framework includes feature extraction using MPSD and a classifier using XGBoost. The MPSD successfully contains meaningful spectral information within the EEG data. As shown in the experiments, the XGBoost successfully detects drowsiness by using spectral information. Our framework shows the most outstanding performance among various feature extraction techniques and machine learning methods, as shown in Section IV-A.

Our framework also has the advantage of being able to obtain the frequency bands and channels that play the most important roles in detecting drowsiness. Through the experiments, we reconfirmed that the alpha and the theta bands are critical to detect drowsiness, which is consistent with the previous works [47], [48]. Furthermore, we have demonstrated that the gamma frequency of $45 - 50\ Hz$ also contributes

to detection of drowsiness, which supports the recent studies on the gamma-based sleepiness detection. Abnormal patterns of gamma waves have been reported to be discovered from patients with neurological disorders (e.g.,/Alzheimer's disease, Parkinson's disease, schizophrenia, and epilepsy) [49].

As a result of our experiment, the Fp1, Fp2, T3, T4, O1, O2, and ECG channels have been shown to play important roles for detecting drowsiness. According to the topological findings, drowsiness detection is possible with the placement of electrodes in the highly relevant regions (frontal, temporal, occipital, and ECG). If only a small number of channels is sufficient for drowsiness detection, we can expect cost and weight reduction of the EEG device. Therefore, we evaluated the performance of various channel combinations. As the number of channels decreases, so does the performance. However, when using the seven channels with the highest importance, performance degradation is less than 2%. In other words, even if only seven channels are used, the performance of drowsiness detection is not dramatically decreased.

To examine the applicability of our framework to drowsiness detection in a mobile device environment, this study acquired additional data through a wireless EEG device using dry sensors. However, the adopted wireless EEG system and R100 were not compatible to be acquired simultaneously due to interference. Since wireless EEG was acquired exclusively during PVT, only approximately 50 minutes of EEG signals per day were used for learning. Detection performance using wireless EEG showed a performance degradation of less than 7% AUC compared to wired EEG, and a value of 0.70 or more in most performance indicators. Two main factors contribute to performance degradation. The first factor is that a relatively small amount of EEG was used for learning, compared to the wired EEG. In the case of the wired EEG, we used five hours per day, whereas only 50 minutes a day of EEG data were available for wireless EEG. Since a small amount of learning data causes performance degradation, it is necessary to accumulate sufficient learning data for wireless brain waves. The second factor is the instability of the dry sensors in the wireless EEG. Although the dry sensors have the advantage of being easy to wear, they have lower signal quality and higher motion-sensitivity than the wet sensors. Nevertheless, performance degradation is not drastic; therefore, if these factors are improved, good and stable detection performance can be expected in the mobile device environment using wireless EEG. In future work, we need to develop and implement an algorithm to handle the EEG artifacts and their effects on the performance to detect drowsiness, especially for dry wireless EEG system.

During the data acquisition, subjects were given various constraints such as long measurements, limited movement, and controlled sleep time. Such constraints make recruiting difficult, which resulted in a small data: eight subjects for the case of our study. Since the amount of data was insufficient, the performances of the deep learning methods such as FCN and CNN were not outstanding. According to recent researches, deep learning has shown a state of the art performance in various fields [69]–[75]. However, deep learning can automatically extract meaningful features and achieve great performance only with a sufficient amount of data. As more data is collected later, we hope not only to achieve better performance, but also to uncover new useful features through deep learning. Another limitation of the proposed framework due to the lack of data is that only the intra-subject approach was considered. One model was trained for one subject, which means that if there is a new subject, the new model will need to be trained. For our model to generalize over any subjects, more data should be accumulated.

ACKNOWLEDGMENT

The authors thank Kiwon Lee, PhD (CEO, YBrain, Inc., Republic of Korea), Jaques Reifman, PhD (Director, Biotechnology High Performance Computing Software Applications Institute, USA) and their team members for supporting the wireless EEG monitoring system and the PC-PVT platform.

REFERENCES

[1] P. M. Krueger and E. M. Friedman, "Sleep duration in the United States: A cross-sectional population-based study," *Amer. J. Epidemiol.*, vol. 169, no. 9, pp. 1052–1063, 2009.

[2] J.-H. Kim, K. R. Kim, K. H. Cho, K.-B. Yoo, J. A. Kwon, and E.-C. Park, "The association between sleep duration and self-rated health in the Korean general population," *J. Clin. Sleep Med.*, vol. 9, no. 10, pp. 1057–1064, Oct. 2013.

[3] J. K. Walsh, C. Coulouvrat, G. Hajak, M. D. Lakoma, M. Petukhova, T. Roth, N. A. Sampson, V. Shahly, A. Shillington, J. J. Stephenson, and R. C. Kessler, "Nighttime insomnia symptoms and perceived health in the America Insomnia Survey (AIS)," *Sleep*, vol. 34, no. 8, pp. 997–1011, 2011.

[4] J. Kim, K. In, J. Kim, S. You, K. Kang, J. Shim, S. Lee, J. Lee, S. Lee, C. Park, and C. Shin, "Prevalence of sleep-disordered breathing in middle-aged Korean men and women," *Amer. J. Respiratory Crit. Care Med.*, vol. 170, no. 10, pp. 1108–1113, Nov. 2004.

[5] Y. Hwangbo, W.-J. Kim, M. K. Chu, C.-H. Yun, and K. I. Yang, "Habitual sleep duration, unmet sleep need, and excessive daytime sleepiness in Korean adults," *J. Clin. Neurol.*, vol. 12, no. 2, pp. 194–200, Apr. 2016.

[6] S. Suh, H.-C. Yang, C. P. Fairholme, H. Kim, R. Manber, and C. Shin, "Who is at risk for having persistent insomnia symptoms? A longitudinal study in the general population in Korea," *Sleep Med.*, vol. 15, no. 2, pp. 180–186, Feb. 2014.

[7] S. Joo, I. Baik, H. Yi, K. Jung, J. Kim, and C. Shin, "Prevalence of excessive daytime sleepiness and associated factors in the adult population of Korea," *Sleep Med.*, vol. 10, no. 2, pp. 182–188, 2009.

[8] S. M. W. Rajaratnam, M. E. Howard, and R. R. Grunstein, "Sleep loss and circadian disruption in shift work: Health burden and management," *Med. J. Aust.*, vol. 199, no. 8, pp. S11–S15, Oct. 2013.

[9] M. Daley, M. Charles Morin, M. LeBlanc, J.-P. Grégoire, and J. Savard, "The economic burden of insomnia: Direct and indirect costs for individuals with insomnia syndrome, insomnia symptoms, and good sleepers," *Sleep*, vol. 32, no. 1, pp. 55–64, 2009.

[10] D. R. Hillman and L. C. Lack, "Public health implications of sleep loss: The community burden," *Med. J. Aust.*, vol. 199, no. 8, pp. S7–S10, Oct. 2013.

[11] C.-H. Yun, H. Kim, S. K. Lee, S. Suh, S. H. Lee, S.-H. Park, R. J. Thomas, R. Au, and C. Shin, "Daytime sleepiness associated with poor sustained attention in middle and late adulthood," *Sleep Med.*, vol. 16, no. 1, pp. 143–151, Jan. 2015.

[12] Y. J. Lee, J. Park, S. Kim, S.-J. Cho, and S. J. Kim, "Academic performance among adolescents with behaviorally induced insufficient sleep syndrome," *J. Clin. Sleep Med.*, vol. 11, no. 1, pp. 61–68, Jan. 2015.

[13] M. L. Zeek, M. J. Savoie, M. Song, L. M. Kennemur, J. Qian, P. W. Jungnickel, and S. C. Westrick, "Sleep duration and academic performance among student pharmacists," *Amer. J. Pharmaceutical Educ.*, vol. 79, no. 5, p. 63, Jun. 2015.
[14] S. Y. Kim, S.-G. Kim, S. Sim, B. Park, and H. G. Choi, "Excessive sleep and lack of sleep are associated with slips and falls in the adult Korean population: A population-based cross-sectional study," *Medicine*, vol. 95, no. 4, p. e2397, Jan. 2016.
[15] J. Connor, R. Norton, S. Ameratunga, E. Robinson, I. Civil, R. Dunn, J. Bailey, and R. Jackson, "Driver sleepiness and risk of serious injury to car occupants: Population based case control study," *Brit. Med. J.*, vol. 324, no. 7346, pp. 1125A–1128A, 2002.
[16] P. M. Forsman, B. J. Vila, R. A. Short, C. G. Mott, and H. P. A. Van Dongen, "Efficient driver drowsiness detection at moderate levels of drowsiness," *Accident Anal. Prevention*, vol. 50, pp. 341–350, Jan. 2013.
[17] D. Sandberg, T. Akerstedt, A. Anund, G. Kecklund, and M. Wahde, "Detecting driver sleepiness using optimized nonlinear combinations of sleepiness indicators," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 1, pp. 97–108, Mar. 2011.
[18] T. H. Chang, C. S. Hsu, C. Wang, and L. K. Yang, "Onboard measurement and warning module for irregular vehicle behavior," *IEEE Trans. Intell. Transp. Syst.*, vol. 9, no. 3, pp. 501–513, Sep. 2008.
[19] S. Ftouni, T. L. Sletten, M. Howard, C. Anderson, M. G. Lenné, S. W. Lockley, and S. M. W. Rajaratnam, "Objective and subjective measures of sleepiness, and their associations with on-road driving events in shift workers," *J. Sleep Res.*, vol. 22, no. 1, pp. 58–69, Feb. 2013.
[20] D. F. Dinges and J. W. Powell, "Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations," *Behav. Res. Methods, Instrum. Comput.*, vol. 17, no. 6, pp. 652–655, Nov. 1985.
[21] D. F. Dinges, N. L. Rogers, J. Dorrian, and C. A. Kushida, "Psychomotor vigilance performance: Neurocognitive assay sensitive to sleep loss," in *Sleep Deprivation*. New York, NY, USA: Marcel Dekker, 2005, pp. 39–70.
[22] J. Lim and D. F. Dinges, "Sleep deprivation and vigilant attention," *Ann. New York Acad. Sci.*, vol. 1129, pp. 305–322, May 2008.
[23] M. Basner, D. Mollicone, and D. F. Dinges, "Validity and sensitivity of a brief psychomotor vigilance test (PVT-B) to total and partial sleep deprivation," *Acta Astron.*, vol. 69, no. 11, pp. 949–959, 2011.
[24] D. F. Dinges, M. M. Mallis, G. Maislin, and J. W. Powell, "Evaluation of techniques for ocular measurement as an index of fatigue and as the basis for alertness management," Nat. Highway Traffic Saf. Admin., Washington, DC, USA, Tech. Rep. DOT HS 808 762. 1998.
[25] M. L. Jackson, G. A. Kennedy, C. Clarke, M. Gullo, P. Swann, L. A. Downey, A. C. Hayley, R. J. Pierce, and M. E. Howard, "The utility of automated measures of ocular metrics for detecting driver drowsiness during extended wakefulness," *Accident Anal. Prevention*, vol. 87, pp. 127–133, Feb. 2016.
[26] E. Aidman, C. Chadunow, K. Johnson, and J. Reece, "Real-time driver drowsiness feedback improves driver alertness and self-reported driving performance," *Accident Anal. Prevention*, vol. 81, pp. 8–13, Aug. 2015.
[27] M. Corbett, "Science & technology watch: A drowsiness detection system for pilots: Optalert," *Aviation, Space, Environ. Med.*, vol. 80, no. 2, p. 149, Feb. 2009.
[28] M. Fabbri, F. Provini, E. Magosso, A. Zaniboni, A. Bisulli, G. Plazzi, M. Ursino, and P. Montagna, "Detection of sleep onset by analysis of slow eye movements: A preliminary study of MSLT recordings," *Sleep Med.*, vol. 10, no. 6, pp. 637–640, 2009.
[29] J. A. Horne and L. A. Reyner, "Counteracting driver sleepiness: Effects of napping, caffeine, and placebo," *Psychophysiology*, vol. 33, no. 3, pp. 306–309, 1996.
[30] S. McGuire, U. Müller, E.-M. Elmenhorst, and M. Basner, "Inter-individual differences in the effects of aircraft noise on sleep fragmentation," *Sleep*, vol. 39, no. 5, pp. 1107–1110, 2016.
[31] L. N. Boyle, J. Tippin, A. Paul, and M. Rizzo, "Driver performance in the moments surrounding a microsleep," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 11, no. 2, pp. 126–136, 2008.
[32] Z. Mardi, S. N. M. Ashtiani, and M. Mikaili, "EEG-based drowsiness detection for safe driving using chaotic features and statistical tests," *J. Med. Signals Sens.*, vol. 1, no. 2, pp. 130–137, Aug. 2011.
[33] M. V. M. Yeo, X. Li, K. Shen, and E. P. V. Wilder-Smith, "Can SVM be used for automatic EEG detection of drowsiness during car driving?" *Saf. Sci.*, vol. 47, no. 1, pp. 115–124, 2009.
[34] A. G. Correa, L. Orosco, and E. Laciar, "Automatic detection of drowsiness in EEG records based on multimodal analysis," *Med. Eng. Phys.*, vol. 36, no. 2, pp. 244–249, 2014.
[35] J. Faber, "Detection of different levels of vigilance by eeg pseudo spectra," *Neural Netw. World*, vol. 14, pp. 285–290, Jan. 2004.
[36] H. Han and K.-Y. Song, "Electroencephalogram-based driver drowsiness detection system using errors-in-variables (EIV) and multilayer perceptron (MLP)," *J. Korean Inst. Commun. Inf. Sci.*, vol. 39, no. 10, pp. 887–895, 2014.
[37] T.-P. Jung, S. Makeig, M. Stensmo, and T. J. Sejnowski, "Estimating alertness from the EEG power spectrum," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 1, pp. 60–69, Jan. 1997.
[38] R. Nikhil Pal, C.-Y. Chuang, L.-W. Ko, C.-F. Chao, T.-P. Jung, S.-F. Liang, and C.-T. Lin, "EEG-based subject- and session-independent drowsiness detection: An unsupervised approach," *EURASIP J. Adv. Signal Process.*, vol. 2008, pp. 192:1–192:11, Jan. 2008.
[39] C. Papadelis, Z. Chen, C. Kourtidou-Papadeli, P. D. Bamidis, I. Chouvarda, E. Bekiaris, and N. Maglaveras, "Monitoring sleepiness with on-board electrophysiological recordings for preventing sleep-deprived traffic accidents," *Clin. Neurophysiol., Off. J. Int. Fed. Clin. Neurophysiol.*, vol. 118, no. 9, pp. 1906–1922, Sep. 2007.
[40] A. Sahayadhas, K. Sundaraj, and M. Murugappan, "Drowsiness detection during different times of day using multiple features," *Australas. Phys. Eng. Sci. Med.*, vol. 36, no. 2, pp. 243–250, Jun. 2013.
[41] A. Subasi, "Automatic recognition of alertness level from EEG by using neural network and wavelet coefficients," *Expert Syst. Appl.*, vol. 28, no. 4, pp. 701–711, 2005.
[42] H.-S. Choi, B. Lee, and S. Yoon, "Biometric authentication using noisy electrocardiograms acquired by mobile sensors," *IEEE Access*, vol. 4, pp. 1266–1273, 2016.
[43] S. Choi, S. Kim, J. Seo, J. Y. Park, and S. Yoon, "Wearable and wireless measurement system for evaluating penile tumescence," in *Proc. IEEE Biomed. Circuits Syst. Conf. (BioCAS)*, Oct. 2015, pp. 1–4.
[44] 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. (BCB)*, 2018, pp. 111–121.
[45] M. Y. Khitrov, S. Laxminarayan, D. Thorsley, S. Ramakrishnan, S. Rajaraman, N. J. Wesensten, and J. Reifman, "PC-PVT: A platform for psychomotor vigilance task testing, analysis, and prediction," *Behav. Res. Methods*, vol. 46, no. 1, pp. 140–147, Mar. 2014.
[46] C. François, J. Wertz, M. Kirkove, and J. G. Verly, "Evaluation of the performance of an experimental somnolence quantification system in terms of reaction times and lapses," in *Proc. 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2014, pp. 5820–5823.
[47] A. A. Putilov and O. G. Donskaya, "Construction and validation of the EEG analogues of the Karolinska sleepiness scale based on the Karolinska drowsiness test," *Clin. Neurophysiol.*, vol. 124, no. 7, pp. 1346–1352, 2013.
[48] C. E. D. Alloway, R. D. Ogilvie, and C. M. Shapiro, "The alpha attenuation test: Assessing excessive daytime sleepiness in narcolepsy-cataplexy," *Sleep*, vol. 20, no. 4, pp. 258–266, 1997.
[49] E. M. Whitham, K. J. Pope, S. P. Fitzgibbon, T. Lewis, R. C. Clark, S. Loveless, M. Broberg, A. Wallace, D. DeLosAngeles, P. Lillie, A. Hardy, R. Fronsko, A. Pulbrook, and J. O. Willoughby, "Scalp electrical recording during paralysis: Quantitative evidence that EEG frequencies above 20 Hz are contaminated by EMG," *Clin. Neurophysiol.*, vol. 118, no. 8, pp. 1877–1888, 2007.
[50] C.-T. Lin, R.-C. Wu, S.-F. Liang, W.-H. Chao, Y.-J. Chen, and T.-P. Jung, "EEG-based drowsiness estimation for safety driving using independent component analysis," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 52, no. 12, pp. 2726–2738, Dec. 2005.
[51] S.-H. Hsu and T.-P. Jung, "Monitoring alert and drowsy states by modeling EEG source nonstationarity," *J. Neural Eng.*, vol. 14, no. 5, 2017, Art. no. 056012.
[52] P.-Y. Tsai, W. Hu, T. B. J. Kuo, and L.-Y. Shyu, "A portable device for real time drowsiness detection using novel active dry electrode system," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Sep. 2009, pp. 3775–3778.
[53] G. Li and W.-Y. Chung, "Estimation of eye closure degree using EEG sensors and its application in driver drowsiness detection," *Sensors*, vol. 14, no. 9, pp. 17491–17515, Sep. 2014.
[54] M. Awais, N. Badruddin, and M. Drieberg, "A hybrid approach to detect driver drowsiness utilizing physiological signals to improve system performance and wearability," *Sensors*, vol. 17, no. 9, p. 1991, 2017.

[55] D. Ribeiro, C. Teixeira, and A. Cardoso, "Eeg-based drowsiness detection platform to compare different methodologies," in *Proc. 4th Exp. Int. Conf.*, Jun. 2017, pp. 318–322.
[56] 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.
[57] D. J. Thomson, "Spectrum estimation and harmonic analysis," *Proc. IEEE*, vol. 70, no. 9, pp. 1055–1096, Sep. 1982.
[58] M. J. Prerau, R. E. Brown, M. T. Bianchi, J. M. Ellenbogen, and P. L. Purdon, "Sleep neurophysiological dynamics through the lens of multitaper spectral analysis," *Physiology*, vol. 32, no. 1, pp. 60–92, 2017.
[59] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, 2016, pp. 785–794.
[60] C. Adam-Bourdarios, G. Cowan, C. Germain-Renaud, I. Guyon, B. Kégl, and D. Rousseau, "The Higgs machine learning challenge," *J. Phys.*, vol. 664, no. 7, 2015, Art. no. 072015.
[61] G. R. Poudel, C. R. H. Innes, P. J. Bones, R. Watts, and R. D. Jones, "Losing the struggle to stay awake: Divergent thalamic and cortical activity during microsleeps," *Hum. Brain Mapping*, vol. 35, no. 1, pp. 257–269, 2014.
[62] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 9, pp. 1263–1284, Sep. 2009.
[63] E. R. DeLong, D. M. DeLong, and D. L. Clarke-Pearson, "Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach," *Biometrics*, vol. 44, pp. 837–845, Sep. 1988.
[64] W. Ting, Y. Guo-Zheng, Y. Bang-Hua, and S. Hong, "EEG feature extraction based on wavelet packet decomposition for brain computer interface," *Measurement*, vol. 41, no. 6, pp. 618–625, 2008.
[65] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
[66] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.
[67] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, p. 436, 2015.
[68] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, no. 1, pp. 321–357, 2002.
[69] S. Min, B. Lee, and S. Yoon, "Deep learning in bioinformatics," *Brief Bioinform.*, vol. 18, no. 5, pp. 851–869, 2016.
[70] J. Baek, B. Lee, S. Kwon, and S. Yoon, "LncRNAnet: Long non-coding RNA identification using deep learning," *Bioinformatics*, vol. 34, no. 22, pp. 3889–3897, 2018.
[71] T. Moon, S. Min, B. Lee, and S. Yoon, "Neural universal discrete denoiser," in *Adv. Neural Inf. Process. Syst. 29*, D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, Eds. Red Hook, NY, USA: Curran Associates, 2016, pp. 4772–4780.
[72] S. Kwon and S. Yoon, "DeepCCI: End-to-end deep learning for chemical-chemical interaction prediction," in *Proc. 8th ACM Int. Conf. Bioinf., Comput. Biol., Health Inform. (ACM-BCB)*, 2017, pp. 203–212.
[73] S. Park, S. Min, H.-S. Choi, and S. Yoon, "Deep recurrent neural network-based identification of precursor microrna," in *Adv. Neural Inf. Process. Syst. 30*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Eds. Red Hook, NY, USA: Curran Associates, 2017, pp. 2891–2900.
[74] B. Lee, J. Baek, S. Park, and S. Yoon, "deepTarget: End-to-end learning framework for microRNA target prediction using deep recurrent neural networks," in *Proc. 7th ACM Int. Conf. Bioinf., Comput. Biol., Health Inform. (BCB)*, 2016, pp. 434–442.
[75] H. K. Kim, S. Min, M. Song, S. Jung, J. W. Choi, Y. Kim, S. Lee, S. Yoon, and H. H. Kim, "Deep learning improves prediction of CRISPR–Cpf1 guide RNA activity," *Nature Biotechnol.*, vol. 36, no. 3, p. 239, 2018.

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

SEONWOO MIN received the B.S. degree in electrical and computer engineering from Seoul National University, Seoul, South Korea, in 2015, where he is currently pursuing the Integrated M.S./Ph.D. degree in electrical and computer engineering. His research interests include machine learning, deep learning, and bioinformatics.

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.

HO BAE received the B.Sc. degree in computer science and the M.Sc. degree in information security from University College London, U.K., in 2007 and 2009, respectively. From 2012 to 2016, he was with the SAP Labs, South Korea. He is currently pursuing the Ph.D. degree in bioinformatics with the Interdisciplinary Program, Seoul National University, South Korea. His research interests include artificial intelligence and bioinformatics.

JEE-EUN YOON received the M.D. degree from the School of Medicine, Chosun University, in 2012. She had completed neurology residency at Soonchunhyang University Seoul Hospital, from 2013 to 2017, and a Sleep Medicine Fellowship at Seoul National University, from 2017 to 2019, South Korea. Her research interests include sleep disorders and sleep apnea pathogenesis.

INHA HWANG received the medical degree from the School of Medicine, Chonbuk National University, Jeonju, South Korea, and she completed a neurology residency from the Veterans Health Service Medical Center, Seoul, South Korea. She is currently doing her Sleep Medicine Fellowship training at Seoul National University Bundang Hospital. Her interests include sleep medicine, cognitive neurology, and neuroelectrophysiology.

DANA OH received the M.D. degree from the School of Medicine, Ewha Womans University, in 2014. She performed a Neurology residency training at Inha University Hospital, Incheon, South Korea. She is currently doing her Sleep Medicine Fellowship training at Seoul National University Bundang Hospital. Her research interests include sleep disorders and sleep apnea pathogenesis.

CHANG-HO YUN received the M.D., M.MSc. and Ph.D. degrees from Seoul National University, in 1994, 1999, and 2003, respectively. He is a Board-Certified Neurologist and a Sleep Specialist, and the Former Research Fellow with the Division of Sleep Medicine, Harvard Medical School, and the Beth Israel Deaconess Medical Center, MA, USA, in 2008 and 2009. He was an Assistant Professor and an Associate Professor of neurology with the College of Medicine, Inha University, from 2003 to 2008. He is currently a Professor of neurology with the Clinical Neuroscience Center, Seoul National University Bundang Hospital, Seongnam, South Korea. His research interests include sleep respiratory and neuro-physiology to elucidate the role of sleep in brain health and cognitive impairment.

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, where he was a Visiting Scholar with the Department of Neurology and Neurological Sciences, from 2016 to 2017. He held research positions at Stanford University and Synopsys, Inc., Mountain View. From 2006 to 2007, he was with Intel Corporation, Santa Clara. He was an Assistant Professor with the School of Electrical Engineering, Korea University, from 2007 to 2012. He is currently a Professor with the Department of Electrical and Computer Engineering, Seoul National University, South Korea. His current research interests include machine learning and artificial intelligence. He was a recipient of the SNU Education Award, in 2018, the IBM Faculty Award, in 2018, the Korean Government Researcher of the Month Award, 2018, the BRIC Best Research of the Year, in 2018, the IMIA Best Paper Award, in 2017, the Microsoft Collaborative Research Grant, in 2017, the SBS Foundation Award, in 2016, the IEEE Young IT Engineer Award, in 2013, and many other prestigious awards.