XGBoost-Based Instantaneous Drowsiness Detection Framework Using Multitaper Spectral Information of Electroencephalography

Hyun-Soo Choi
Seoul National University
Seoul, Korea

Siwon Kim
Seoul National University
Seoul, Korea

Jung Eun Oh
Seoul National University Bundang Hospital
Seongnam, Korea

Jee Eun Yoon
Seoul National University Bundang Hospital
Seongnam, Korea

Jung Ah Park
Seoul National University Bundang Hospital
Seongnam, Korea

Chang-Ho Yun*
Seoul National University Bundang Hospital
Seongnam, Korea
ych333@snubh.org

Sungroh Yoon†
Seoul National University
Seoul, Korea
sryoon@snu.ac.kr

ABSTRACT

The socioeconomic losses caused by extreme daytime drowsiness are enormous in these days. Hence, building a virtuous cycle system is necessary to improve work efficiency and safety by monitoring instantaneous drowsiness that can be used in any environment. In this paper, we propose a novel framework to detect extreme drowsiness using a short time segment (~ 2 s) of EEG which well represents immediate activity changes depending on a person's arousal, drowsiness, and sleep state. To develop the framework, we use multitaper power spectral density (MPSD) for feature extraction along with extreme gradient boosting (XGBoost) as a machine learning classifier. In addition, we suggest a novel drowsiness labeling method by combining the advantages of the psychomotor vigilance task and the electrooculography technique. By experimental evaluation, we show that the adopted MPSD and XGB techniques outperform other techniques used in previous studies. Finally, we identify that spectral components (θ, α, and γ) and channels (Fp1, Fp2, T3, T4, O1, and O2) play an important role in our drowsiness detection framework, which could be extended to mobile devices.

*to whom correspondence should be addressed about clinical opinion
†to whom correspondence should be addressed about machine learning techniques

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
ACM-BCB'18, August 29–September 1, 2018, Washington, DC, USA
© 2018 Association for Computing Machinery.
ACM ISBN 978-1-4503-5794-4/18/08…$15.00
https://doi.org/10.1145/3233547.3233567

KEYWORDS

Electroencephalography; Drowsiness; Alertness; XGBoost; Multitaper Power Spectral Density

ACM Reference Format:
Hyun-Soo Choi, Siwon Kim, Jung Eun Oh, Jee Eun Yoon, Jung Ah Park, Chang-Ho Yun, and Sungroh Yoon. 2018. XGBoost-Based Instantaneous Drowsiness Detection Framework Using Multitaper Spectral Information of Electroencephalography. In *ACM-BCB '18: 9th ACM Int'l Conf. on Bioinformatics, Computational Biology, and Health Informatics, Aug. 29–Sept. 1, 2018, Washington, DC, USA.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3233547.3233567

1 INTRODUCTION

Proper amount and quality of sleep are of paramount importance to humans for performing daytime tasks and maintaining cognitive function. However, most people are chronically sleep-deprived due to social requirements or personal preferences and suffer from major sleep disorders [1–4]. More than 30% of adults sleep less than 7 h per day [5], and sleep apnea and insomnia are very common [4, 6]. Over 10% of adults suffer from extreme daytime drowsiness [5, 7]. Lowered attention associated with daytime drowsiness hinders work productivity and increases the risk of accidents [8]. The socioeconomic burden imposed by drowsiness is considerable [9–11] and includes low academic achievement [12, 13] and increased risk of falling and traffic or workplace accident [14, 15]. Hence, it is necessary to build a virtuous cycle system to improve work efficiency and safety by monitoring a person's drowsiness and providing appropriate feedback [16].

The existing techniques for drowsiness monitoring can be divided into two categories: task performance measurement and biosignal-based detection. The former includes vehicle motion monitoring [17–20] and the psychomotor vigilance task (PVT)[21–24]. Vehicle motion monitoring has been developed for restrictive use

during driving. The PVT has a limitation that continuous monitoring is impossible since a subject is required to stop the ongoing task and respond to given stimuli for 3 or 10 min [21, 24]. One of the methods that use biosignals is electrooculography (EOG) monitoring [20, 25–29]. The EOG-based technique detects drowsiness by analyzing blinking (frequency, speed), and eye movements by means of a video camera or an infrared monitoring device. Since the EOG-based technique measures the secondary change caused by the increase of homeostatic sleep load in the nervous system, it has a low temporal resolution. Further, its results are affected by the environment (wind, temperature, humidity) and personal habits.

In our work, we adopted electroencephalography (EEG) to monitor instantaneous drowsiness. EEG that measures electrical activity in the brain is the standard technique used in sleep research. EEG is useful to discriminate the state of alertness, drowsiness, or sleep, by detecting instantaneous wave changes specific to each state. Prior studies have confirmed the presence of EEG frequency change specific to drowsiness [30–32]. Some machine learning-based approaches have also been adopted [33–44]. However, in these studies, the definition of drowsy state has not been consistent. Moreover, since most of the algorithms use a long time segment of EEG (> 5 s), they are not suitable for instantaneous drowsiness detection (refer to Section 2.3).

In this work, we propose a novel framework to detect extreme drowsiness using a short time segment (~ 2 s) of EEG. First, to label drowsiness for machine learning purposes, we combine the advantages of PVT and EOG techniques. The EOG outputs task-independent measures of sleepiness over a prolonged period. Applying the PVT as a standard reference of extreme sleepiness to a long recording of EOG, we suggest a novel method to define the drowsiness label in an ensemble manner. In addition, we use multitaper power spectral density (MPSD) for feature extraction from the short segment of EEG, and extreme gradient boosting (XGBoost) as a machine learning classifier. In the experiment, we verify the validity of the adopted MPSD and XGBoost techniques by illustrating that our techniques outperform the other techniques (based on band power, wavelet decomposition, single taper power spectral density). Finally, we analyze the extent of contribution of each spectral feature component to drowsiness detection. This analysis is conducted by using the branch gain of XGBoost and suggests the important frequency bands and channels that will be useful in drowsiness detection.

2 RELATED WORKS

2.1 Task Performance Evaluation-based Estimation of Drowsiness

Examples of task performance evaluations are the vehicle motion monitoring module [17–20] and the PVT [21–23]. The vehicle monitoring module is based on changes in the lane keeping state and the driving direction. The module is vulnerable to variations in vehicle type, road driving condition, weather, and so on. The PVT measures subjects' response times to randomly presented visual stimuli [21, 24]. For the 10 min in a standard PVT session, the subjects are instructed to press a button with the thumb of the dominant hand as soon as a visual stimulus appears on the machine screen. The inter-stimulus interval is random, ranging from 2 to 10 s. Response times are measured in milliseconds for all the stimuli in each trial. The response times are a validated indicator of the alertness level. The PVT is known to accurately reflect arousal changes arising from sleep deprivation and sleeping disorders [21–23]. The task measures not only the level of vigilant attention but also the influence of state instability. According to the state instability theory, sleep deprivation leads to a fluctuation in sustained attention because of the interaction of involuntary sleep-initiating and counteracting wake-maintaining systems, and eventually to increases in the lapses (errors of omission). However, to measure the arousal with the PVT, it is necessary to stop the original task and perform the task for a certain time (3 min or 10 min), and this interrupts the work continuity. In our study, we use PC-PVT (Biotechnology High Performance Computing Software Applications Institute, http://bhsai.org/software/pcpvt, MD, USA) [45], which is functionally equivalent to PVT except for not using additional hardware but using a personal computer. The arousal information obtained by the PC-PVT platform is used for labeling drowsiness for an EEG segment to train the XGBoost-based drowsiness detection model.

2.2 EOG-based Estimation of Drowsiness

EOG monitoring analyzes eye blinking (frequency, speed, eye winding) and movements using a video camera or an infrared surveillance device. Daytime drowsiness is closely related with the increased frequency of eye blinking, decreased speed, slow eye movements, and increased eye closure time [20, 25, 27–29]. EOG can be measured with relatively simple equipment and is advantageous in that EOG can be used in any action and is not limited to a specific task. However, EOG has a low temporal resolution and can be affected by ambient lighting, humidity, wind, and personal habits. One of the recently developed EOG-based drowsiness estimators is R100 [46]. The R100 measures the degree of drowsiness by sensing various types of eye movements using glasses equipped with a high-speed camera. It has the advantage of being able to measure for a long time in various work environments. However, detecting instantaneous drowsiness with the R100 is difficult because it requires the previous 60-s data of eye movement to estimate drowsiness in a stable manner. In this study, we use the R100 device for allocating a label (normal or drowsy) to an EEG segment for XGBoost training, in which the labeling is carried out by using the drowsiness level measured by the R100 device for a long time EEG segment in a free daily situation.

2.3 EEG-based Estimation of Drowsiness

EEG is a standard technique of sleep research that measures the electrical activity in the brain. EEG shows immediate waveform changes depending on the arousal, drowsiness, and sleep states. Putilov *et al.* [47] reported that in the drowsy state, the power increases in the alpha band and the theta band when the eyes of the subject are open, and the theta band power increases only when the subject closes the eyes. Alloway *et al.* [48] reported that, in the drowsy state, the power in the alpha band decreases when the subject's eyes are closed and increases when the eyes are open.

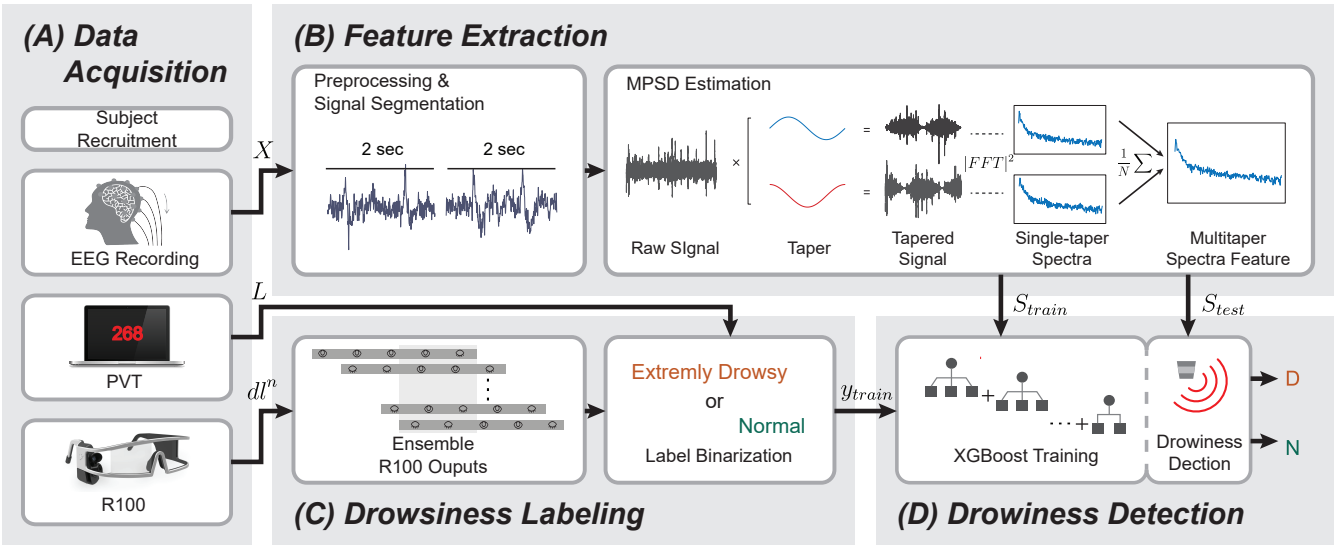


Figure 1: Overall scheme of proposed framework

To train the EEG-based drowsiness detection model, the previous study used various drowsiness label definition methods that can be categorized into three types. The first type is to ask the subjects about their subjective drowsiness. For this, we can use a questionnaire such as the Epworth sleepiness scale [49] and Karolinask sleepiness scale [47]. Although the label of drowsiness can be easily obtained, its validity is hard to verify due to personal subjectivity. The second type uses the task performance evaluations, which is described in Section 2.1. This type of labeling methods use automotive detection modules [37, 38, 42, 50, 51] or response speeds to specific stimuli similar to the PVT [36, 51, 52]. The third type uses the EOG-based method. In Gang's work [53], the label is defined using EOG, but only one parameter (degree of eyelid closure) is considered. Our proposed drowsiness labeling method uses both R100 and PVT, where R100 gives a task-independent measure of the sleepiness level, and PVT provides a standard reference of extreme sleepiness labeling.

2.4 Feature Extraction for EEG Signals

Various feature extraction techniques have been developed to obtain meaningful information from EEG signals and applied to drowsiness detection. Studies have been carried out using a specific frequency band power (BP), which is clinically easy to interpret, such as θ: $4-8$ Hz, α: $8-16$ Hz [36, 38, 54, 55]. Some studies have used power spectral density (PSD) estimation or wavelet decomposition to obtain more detailed frequency information [41, 50, 52]. There also have been studies that have developed and applied feature extraction based on information theory [37, 55].

Unlike in the BP-based technique, the PSD-based technique calculates the power for the detailed frequency range. However, the conventional PSD has the issues of high bias and high variance in estimation. To overcome the high-bias issue, a single-taper PSD (SPSD) technique was developed [56]. However, the SPSD technique cannot solve the high-variance problem of estimation. To solve this, a multitaper power spectral density (MPSD) [57] based method has been developed. MPSD integrates several independent SPSDs to facilitate precise estimation. In Preraue's research [58], MPSD was more discriminative than SPSD in EEG when the stages of the sleep cycle changed. In our study, we use MPSD to extract spectral features for a small segment of EEG.

2.5 Machine Learning Methods for Drowsiness Detection

Several machine learning techniques have been used to detect drowsiness. In the study conducted by Jung [36], the machine learning approach, linear regression, was first applied. Since then, artificial neural networks (ANNs) have been used in several studies [34, 41, 51, 52]. Recently, support vector machines (SVMs) have been widely used owing to its structural efficiency [42, 51, 54]. In our study, we use the extreme gradient boosting (XGBoost) method [59]. XGBoost is an efficient and scalable variant of the gradient boosting machine. Since XGBoost has several advantages in terms of ease of use, accuracy, and scalability, XGBoost has won the recent machine learning competitions [60]. Furthermore, XGBoost can estimate the importance of each input feature from the branch gain. Using this branch gain in XGBoost, we can analyze the importance of each spectral feature used in our study, making it easy for us to understand which frequency bands or channels play a significant role in drowsiness detection.

3 METHODS

As depicted in Figure 1, the overall scheme of the proposed framework is composed of four steps: data acquisition, feature extraction, drowsiness labeling, and drowsiness detection. In the data acquisition step, we acquired EEG signals from recruited subjects . In addition, to support label formulation, we conducted photo-oculography measurement (R100), and task performance measurement (PC-PVT) (details are presented in Section 3.1). In the feature extraction step,

Table 1: Total sleep time (minutes) according to sleep condition

Non sleep-deprived		Sleep-deprived	
Day -6 ~ -2	Day -1	Day -6 ~ -2	Day -1
441.25±38.50	431.25±31.82*	440.52±44.43	216.25±30.21*

*Total sleep time of day -1 on each condition showed significant different ($p < 0.05$).

we extracted the MPSD feature [57] from EEG segments obtained by preprocessing (see Section 3.2). In the drowsiness labeling step, lapse-based drowsiness labeling is provided for each EEG segment by using ensemble R100 outputs and lapses obtained from PC-PVT (see Section 3.3). In the final step, XGBoost [59] is trained with the input feature S_{train} from the feature extraction step and its label y_{train} from the label formulation step. The trained XGBoost predicts the drowsiness corresponding to a new EEG segment (see Section 3.4).

3.1 Data Acquisitions

3.1.1 Subjects. We recruited eight healthy subjects (four male and four female with aged 20 – 40 years) who were nonsmokers and free from sleep disorders such as obstructive sleep apnea, insomnia, restless legs syndrome, and primary hypersomnia. None were taking any medication that affects sleep or alertness, and they reported no engagement in shiftwork for the preceding one-year period and no travel across the two-time zones during the one month prior to the enrollment. All the participants were regular sleepers, sleeping 6 – 8 h daily, in an average of intermediate chorotype.

The institutional review board of Seoul National University Bundang Hospital approved the study protocol. All the participants provided written informed consent.

3.1.2 Experimental Protocol. The schematic diagram of the data acquisition protocol in this study is depicted in Figure 2. Participants were required to sleep regularly for at least 7 h for the entire study period except for the D-1 night for the sleep-deprived condition when the subjects should sleep less than four hours. The subjects' compliance to the instruction was monitored through a sleep diary and was generally accepted as shown in Table 1. All the subjects slept less than 4 h on the single night to ensure sleep deprivation. One had average 6.0 h of sleep for the non sleep-deprived conditions, and the other three had insufficient sleep from D-6 to D-2 and were thus sleep deprived (6.1, 6.9, and 6.8 h of sleep, respectively). We did not exclude these subjects although it may attenuate or accentuate the contrast between sufficient-sleep and sleep-deprived conditions. We measured EEG, R100, and PVT signals in an isolated space during the day (9:30 to 18:30 hours) for the subjects who slept in two different night-time sleeping conditions (see Table 1) in their residences. The monitoring interval for each sleep condition should be at least two weeks, and each condition should be randomly assigned. Subjects are prohibited from drinking caffeine and alcohol for 48 h before the measurement. After measurement in the sleep-deprived condition, the subjects should sleep more than 8 h at nighttime on at least three days. On the days of measurement, the

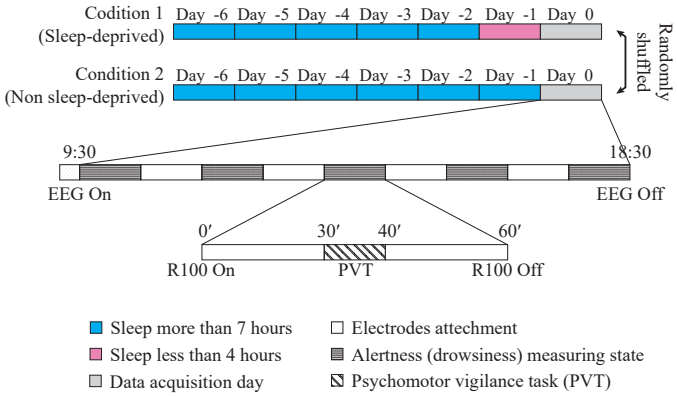


Figure 2: Illustration of data acquisition protocol

subjects had breakfast before starting the measurement and lunch between 12:30 pm and 1:30 pm. The subjects were allowed to drink water, but prohibited from caffeinated drinks, alcohol, and smoking.

3.1.3 Measurement of EEG signal and alertness (drowsiness). In our study, a wired wet electrode EEG (Beehive Horizon, Grass Technologies, Natus, USA; sampling rate) device was used, and the EEG nodes were attached to a total of 19 channels according to the standard 10-20 system as shown below:

$$\text{set } C = \{FP1, FP2, F7, F3, Fz, F4, F8, T3, T5, T4, T6, C3, Cz, C4, P3, Pz, P4, O1, O2\}. \quad (1)$$

For the data acquisition, we adopted a referential montage where all the electrodes were referenced to left mastoid electrode (A1).

EEG measurements were performed in a separate laboratory from 9:30 am to 6:30 pm. Naps during the measurements were prohibited and monitored through real-time video and EEG monitoring. Per each sleeping condition, a total of five alertness measurements were performed at intervals of 2 h. Each measurement was 60-min long and consisted of 60-min R100 recording and 10-min PVT in the background of continuous EEG recording. The PVT was performed in the middle of R100 tracking, started after 30 minutes with the subjects wearing R100. To synchronize the PVT response with the EEG, PC-PVT was installed in the laptop connected to the EEG device. There is no restriction of behavior except during PVT.

3.2 Feature Extraction from EEG

3.2.1 EEG Signal Preprocessing. The signal $\mathbf{x}^j$ from the jth channel was obtained by the difference between the electric potential $\mathbf{i}^j$ obtained from the corresponding channel and the $\mathbf{i}^{A1}$ signal of the A1 node attached to the left mastoid.

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

The frequency filter settings were as follows: 1 Hz low-pass filter, 50 Hz high-pass filter, and the sampling rate was 200 Hz. We applied average referential montage and longitudinal bipolar montage, but the other montages did not show any significant difference with the right mastoidal reference in performance.

As mentioned in Section 1, to prevent accidents caused by drowsiness, detecting instantaneous drowsiness by using a short time

signal is necessary. In our approach, the EEG is divided into specific window sizes without overlap. When the window size is defined as w, the EEG segment of the ith window X_i of the jth channel $\mathbf{x}_i^j$ can be expressed as below:

$$\mathbf{x}_i^j = [x^j_{(i-1)*w+1}, x^j_{(i-1)*w+2}, ...x^j_{i*w}]. \quad (3)$$

The generated EEG segments are used as the input to the machine learning classifier.

3.2.2 Multitaper Spectral Feature Extraction. As described in Section 2.4, in this study, feature extraction is performed on EEG signals using multitaper power spectral density (MPSD) [57]. For the signal $\mathbf{x}_i^j$ in (3) and given a set of L tapers, the lth single taper PSD $s_i^{j(l)}(f)$ [56] at frequency f is calculated as

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

where $w_k^{(l)}$ is the value of the lth taper function at time k and Δt indicates the sampling duration in each EEG segment. Then, MPSD at frequency f is obtained by

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

Equation (5) can be used to compute the spectral information for the m frequencies included in the EEG segment. Thus, the MPSD value of the EEG segment $\mathbf{x}_i^j$ is given as below:

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

A feature vector $\mathbf{S}_i = [\mathbf{s}_i^1, \mathbf{s}_i^2, ..., \mathbf{s}_i^n] \in \mathrm{R}^{m*n}$ can be generated by concatenating n number of s_i^j from the EEG segment X_i.

3.3 Drowsiness Labeling

3.3.1 R100-based Drowsiness Labeling. For the ith EEG segment feature S_i obtained from the EEG feature extraction step, it is necessary to define a label indicating drowsiness to train the classifier. In this paper, we define the label of an EEG segment by using the R100 output that indicates the level (degree) of drowsiness. Since the EEG segment is relatively short compared to the time required by R100, the output dl^n of R100 overlaps the EEG segment duration, as seen in Figure 3. Hence, we can provide the drowsiness level of an EEG segment using one of the R100 outputs overlapped with the EEG segment. On the other hand, in our work, to achieve robustness, we adopt an ensemble approach that combines all overlapped R100 outputs as shown in Figure 3. The average of all R100 outputs dl_i^n overlapping the ith EEG segment is defined as the ensemble drowsiness level dl_i^{ens} as follows:

$$dl_i^{ens} = \frac{1}{N} \sum_{n=1}^{N} dl_i^n. \quad (7)$$

To give the justification of the proposed average ensemble method (AE), two additional methods were implemented and empirically compared in the experiment. The first method used the majority among several dls and the second to used the first output dl^1. The methods are called the majority ensemble (ME) and the first output (FO).

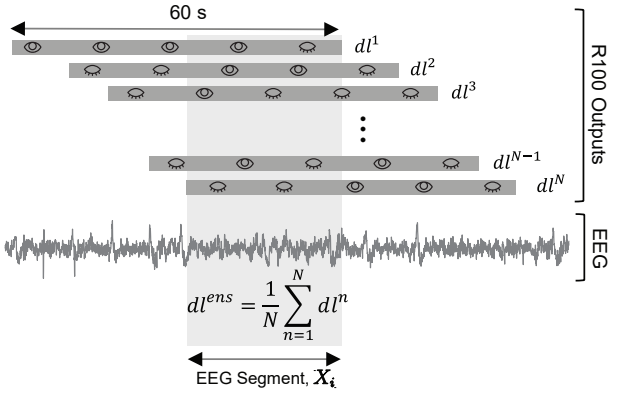


Figure 3: Definition of drowsiness level for EEG segment by ensembling multiple R100 outputs

To convert the dl_i^{ens} value to a binary classification label, the following formula is defined:

$$y_i = \begin{cases} 1 & \text{if } dl_i^{ens} \geq threshold \\ 0 & \text{if } dl_i^{ens} < threshold. \end{cases} \quad (8)$$

$y_i = 1$ represents the extreme drowsiness state and $y_i = 0$ represents the normal state. The next section describes how to select the threshold.

3.3.2 Threshold searching. Lapse refers to the case in which a subject has a response time (RT) greater than or equal to a certain time (t) during PVT. The occurrence of lapse implies that the subject is in an extremely drowsy state. If the PVT response time in X_i is RT_i, the occurrence of lapse L_i is defined as follows:

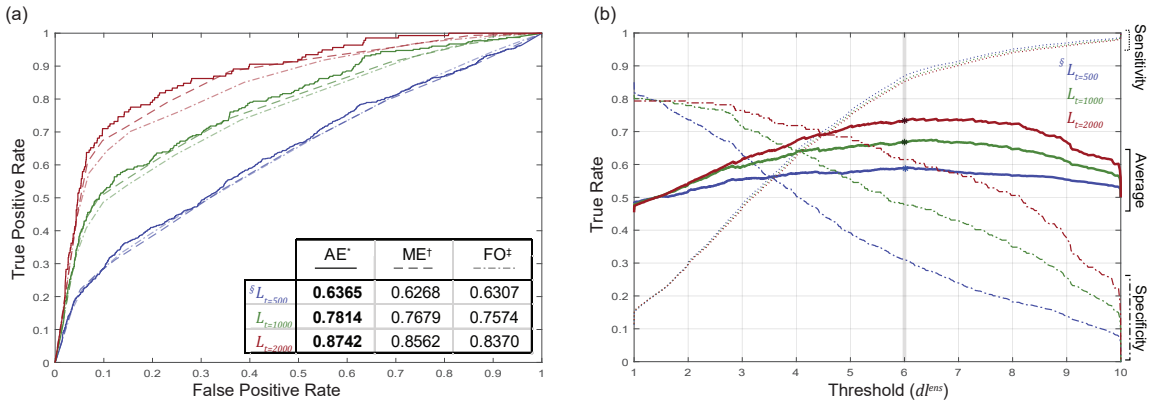
$$L_i = \begin{cases} 1 & \text{if } RT_i \geq t \\ 0 & \text{if } RT_i < t \end{cases}. \quad (9)$$

To determine a meaningful threshold of (8) using L_i obtained from PVT, we use Francois' method [46]. For a specific t, we can set the *threshold* in (8) so that y_i in (8) becomes similar to L_i in (9) for all EEG segments.

For a specific value of *threshold*, the sensitivity and specificity can be calculated using the difference between L_i and y_i. Therefore, we can plot the receiver operating characteristic (ROC) curve by changing the threshold value for a specific value t. Since sensitivity and specificity have a relation of trade-off, the *threshold* is determined as the value at the point maximizing the mean value of sensitivity and specificity. In our work, we applied this procedure for $t \in \{500, 1000, 2000\}$ and the resulting *threshold* was the same regardless of the value of t, as shown in Figure 4 (b) of the Experimental Section. By applying this threshold value to (8), we can determine y_i for the signal over the entire time the subjects wear R100.

3.4 Drowsiness Detection

We train the XGBoost-based binary classifier for each subject separately with the feature vector S_i defined as the input and the drowsiness label y_i as the output. XGBoost ensembles several classification and regression trees (CART) and outputs the prediction



*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 AUROC 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.

value $\hat{y}_i$ as follows:

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

where f_t represents the tth tree and T is the number of trees. The objective function for training of XGBoost is given by

$$Obj = \sum_{i}^{M} l(y_i, \hat{y}_i) + \sum_{t}^{T} \Omega(f_t), \tag{11}$$

where M is the number of training samples, $l(\cdot)$ denotes a loss function, and $\Omega(\cdot)$ refers to the complexity of trees. We use cross-entropy as the loss function like as given below:

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

In this study, the ratio of the negative label samples to the positive label samples is approximately 8 : 2 because the positive samples indicating the extreme drowsiness state are relatively rare compared to the negative samples indicating normal state. To solve this problem, in calculating $l(y_i, \hat{y}_i)$, the cost-sensitive loss is defined as (13) by multiplying a weight with the positive target.

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

where w_c = (# of positive)/(# of negative).

XGBoost sequentially adds trees to minimize the loss in (11) according to an additive training technique as

$$\hat{y}_{(t)} = \hat{y}_{(t-1)} + f_t(S), \ t = 1, ..., T; \ \hat{y}_0 = 0. \tag{14}$$

The maximum number of trees is limited to 5,000. To prevent overfitting, the area under curve (AUC) value of the validation set is monitored during the learning process and the early stopping technique is used to stop learning when there is no change more than 300 times. In our work, we used 80% of randomly selected data from each subject as train data, 10% for validation and early stopping, and 10% for the test set.

In [61], it has been proven that the object function in (11) converges to the optimal value Obj^* given by

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

where k indicates the index of each leaf, and K is the number of leaves in a tree; λ and γ are the design parameters for balance between the complexity and accuracy of leaf scores. Here,

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

where I_k denotes an indicator set representing the samples arriving at the kth leaf in each tree. The value of Obj^* in (15) measures how good a tree structure is. In practice, we optimize the tree structure by expanding the leaves by using the information gain (*Gain*) (17) during the process of dividing branches in each tree without optimizing the structure considering all the trees. The *Gain* is defined by

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

where the first term denotes the left children score, the second term refers to the right children score, and the third term is the score when no split is done. That is, we can create a tree so that the total gain for the tree is maximized. To this end, if the gain for each leaf is smaller than γ, we do not add an additional branch. This is a kind of pruning techniques in tree-based models.

In addition, since one feature is matched to one branch of the tree, we can compute the sum of leaf gains in the path of the branch for a specific feature. The sum of leaf gains represents the importance for an input feature. In our study, since spectral information was used as a feature, the importance of each frequency channel can be evaluated from the *Gain* values. That is, we can see easily which frequency and which channel play a significant role in our drowsiness detection method. Refer to [59] for details on XGBoost.

Table 2: Subject-specific performance on test data

Subject	Accuracy	AUROC	Sensitivity	Specificity
1	0.7964	0.8871	0.7962	0.7970
2	0.7706	0.8459	0.7707	0.7699
3	0.7718	0.8573	0.7719	0.7710
4	0.8107	0.8891	0.8108	0.8102
5	0.7497	0.8310	0.7501	0.7459
6	0.8163	0.8963	0.8163	0.8160
7	0.7790	0.8681	0.7792	0.7761
8	0.8119	0.8946	0.8119	0.8114
Mean	0.7883	0.8712	0.7884	0.7872
Std	0.0225	0.0229	0.0225	0.0235

4 RESULTS

4.1 Drowsiness Label Evaluation

We have defined the drowsiness label for the EEG segment using R100 and PVT as explained in Section 3.3. We compared the proposed average ensemble (AE) with the other two methods (ME and FO). As described in Section 3.3.2, Figure 4 (a) depicts ROC curves for (9) using the dl^{ens} values obtained by the three methods for label definition. As a result, we can confirm that the proposed AE method shows high AUC for all t; thus, it can be seen that the label defined by the AE method is most related to the lapse during PVT.

Figure 4 (b) shows the validity of the method to determine the $threshold$ described in Section 3.3.2. The sensitivity and specificity between L_i in (9) and y_i in (7) are shown for three lapses depending on t. The bold line in Figure 4 represents the mean value of sensitivity and specificity. We determined the threshold at the point where the mean value becomes maximum. For all t, the mean value of sensitivity and specificity had the maximum value when $dl_{ens} = 6$. In other words, if $dl_{ens} \geq 6$ (=$threshold$), the subjects could be defined as extremely drowsy.

4.2 Individual Drowsiness Detection Performance

To evaluate the performance, we used the following four metrics[1]:

- AUROC: the area under the ROC curve which plots sensitivity (t) versus 1−sensitivity(t) with varying threshold t
- Accuracy = $(TP + TN)/(TP + TN + FP + FN)$
- Sensitivity = $TP/(TP + FN)$
- Specificity = $TN/(TN + FP)$

We calculated the four metrics on the test data for each subject, and the mean and standard deviation over the subjects were also given. The performance of the proposed framework is shown in Table 2.

All the metrics have standard deviations of about 0.022. This means that the proposed drowsiness detection method is robust against variations of EEG signals depending on subjects. Furthermore, most metrics show a reasonable performance above 0.78, implying that the proposed method can be applied to actual applications.

[1] TP: true positive, FP: false positive, FN: false negative, TN: true negative

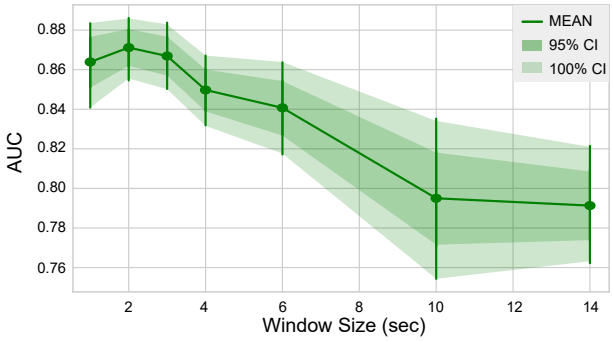
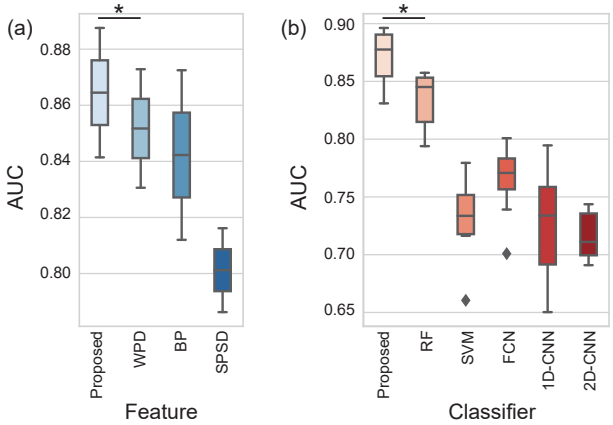


Figure 5: AUC trends depending on window size, w in (3)



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

Figure 6: Comparative Analysis on other (a) features and (b) classifiers

4.3 Validation for Detecting Instantaneous Drowsiness

For the instantaneous detection of extreme drowsiness, the performance for a short EEG segment should be guaranteed. To verify this, we investigated the AUC trends by varying the window size, w in (3). As shown in Figure 5, the proposed framework shows the best accuracy at the window size of 2 s which is small enough to meet the instantaneous detection requirement. Figure 5 shows that a long EEG segment degrades the accuracy, presenting the possibility that a long EEG segment may contain mixed information on normal and drowsiness states.

4.4 Comparative Analysis

To verify the techniques applied to the proposed framework, we compared our techniques with other techniques. Three feature extraction methods and five classifiers were compared. AUC values are shown in Figure 6, and Delong's method [62] was used to confirm the statistical significance between AUC values.

4.4.1 Comparison with Other Feature Extraction Methods. Three feature extraction techniques were compared. The first technique

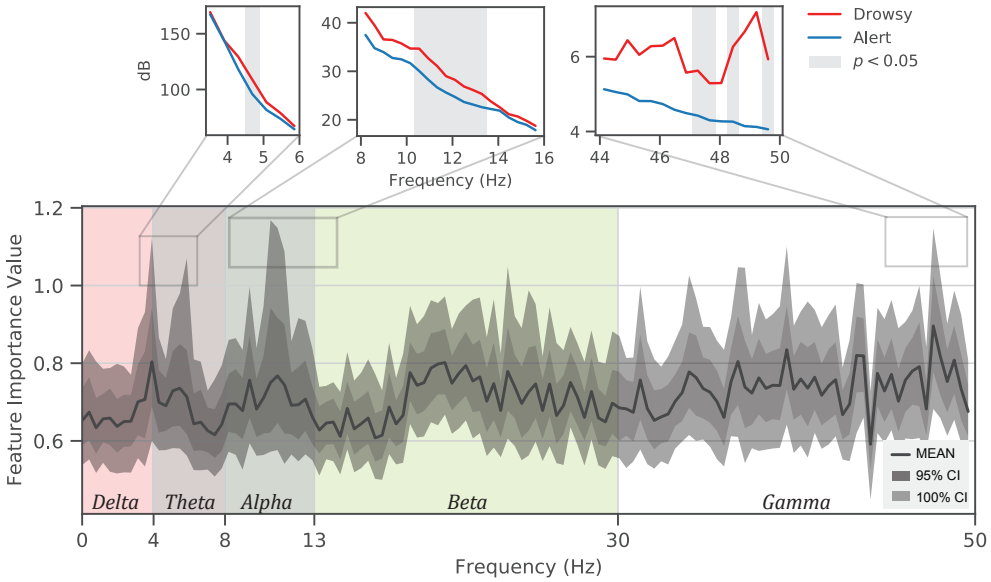


Figure 7: Feature importance as a function of frequency

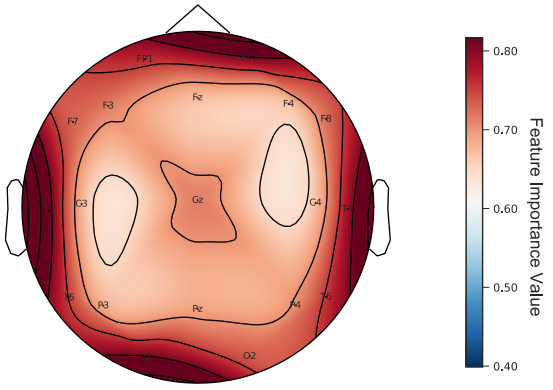


Figure 8: Topographic mapping of feature importance value

[54] uses the EEG powers for 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 technique is the wavelet packet decomposition (WPD) [63], which is known to be suitable for nonstationary signals. The third technique is SPSD based on Welch's method [56]. For the all feature extraction techniques, XGBoost with the same hyper-parameters was used. The comparison result is shown in Figure 6 (a). While SPSD shows the lowest performance, MPSD shows the best performance. This means that the multi-taper technique plays a significant role in representing power spectral information accurately. The statistical test also shows that MPSD achieves a significant improvement of performance.

4.4.2 Comparison with Other Classifiers. To show the outperformance of XGBoost, it was compared with three types of classification algorithms. The first type is random forest (RF) [64] based on the bagging technique of ensemble algorithm series. The second type is SVM [65], and the third type is a deep learning-based model [66]. Since deep learning could be implemented into various architectures, we implemented three architectures: fully connected network (FCN), 1D convolutional network (1D-CNN), and 2D convolutional network (2D-CNN). It was difficult to train the SVM and the deep learning-based model owing to class imbalance. For this reason, the synthetic minority over-sampling technique (SMOTE) [67] was applied additionally to create virtual data for minor classes for both algorithms. The hyper-parameters of each algorithm were found by greedy search.

The comparison result is shown in Figure 6 (b). XGBoost shows the best performance. RF-based detection shows relatively good performance compared to others, whereas SVM-based detection shows relatively low performance as it focuses on major class (normal state). Since our data is not enough to train the deep learning networks, they also do not show good performance even with applied various techniques for enhancing performance, such as drop-out and batch normalization, etc.

4.5 Feature Importance

As described in Section 3.4, we can evaluate the importance of a spectral feature using the gain value obtained from Figure 7, which shows the feature importance trend as a function of frequency. The gain peaks on CI in the figure imply that the frequency channels in the theta band and the alpha band play an important role in drowsiness detection, and this consistent with the fact presented in the existing studies such as the Karolinska drowsiness test [47] and the alpha attenuation test [48]. In addition, we found a novel fact that the range of 45 – 50 Hz in the gamma band takes an important role. Figure 8 is a topology map of importance for each channel.

Fp1, Fp2, T3, T4, O1, and O2 have relatively higher importance than other channels.

5 DISCUSSION

In this study, we have proposed an instantaneous drowsiness detection method that is not restricted to any given situation. Once the training of the detection model is completed, there is no need to perform troublesome tasks such as the PVT. The proposed framework has a great advantage of being able to detect drowsiness during any task such as driving or studying. In addition, as we discussed in Section 4.3, since our method requires only 2 s of the EEG segment for drowsiness detection, there is a great merit that the instantaneous drowsiness state such as lapse can be immediately detected and informed.

The proposed framework uses MPSD for feature extraction and XGBoost for classification learning. The MPSD well represents the spectral information of the EEG segment. Using the spectral information, XGBoost can perform successful detection as shown in the experiments. As discussed in Section 4.1, the techniques adopted in our framework show better performance than the other techniques.

Furthermore, we could identify which the frequency bands and channels that play a major role in drowsiness detection. From the experiment result, we could confirm that the theta band and the alpha band play a major role, as presented in the existing works [47, 48]. We newly found that the $45 - 50Hz$ range of the gamma band plays a major role as well. Irregular gamma waves have been found in the EEGs of patients with neurological disorders (eg, Alzheimer's disease, Parkinson's disease, schizophrenia, and epilepsy) [49]. On the other hand, there is an opinion that EEG-measured gamma waves are meaningless for actual analysis , since gamma waves are easily affected by the artifact of electromyographic activity [49]. There is a need to consider the effectiveness of the gamma wave in detecting actual drowsiness. Our research group will continue to collaborate for a follow-up study.

In our experiment, the Fp1, Fp2, T3, T4, O1, and O2 channels have been evaluated as the most important channels for drowsiness detection. Based on the topological results, drowsiness can be detected by placing the electrodes in the regions that showed high relevance (frontal, temporal, and occipital). If drowsiness can be detected by using only a small number of channels, the computational cost can be reduced. In addition, by reducing the number of sensors in the EEG device, the weight of the drowsiness detection device can be reduced. Achieving this will be a great basis for our researchers' plan to extend the framework to mobile devices in the future.

The limitation of our study is the small amount of data because the number of recruited subjects was only eight. It was difficult to recruit sufficient subjects due to various limitations such as long measurement time, restriction of movement arising from device mounting, and control of sleep time. Due to the limitation of data, deep learning methods such as FCN and CNN showed low performance. However, in recent years, deep learning has become a state of art in many fields [68–74]. If the data is sufficient, deep learning techniques can achieve high classification performance and can automatically learn meaningful features [66]. If more data is accumulated later, we expect to be able to not only achieve high performance but also discover new meaningful features using deep learning.

Another limitation is that proposed framework is limited to the intra-subject approach. In other words, a new learning process is needed for new subjects because a model is constructed only for one subject. To establish a model that can be applied to all subjects considering inter-subject access, data for many subjects should be accumulated.

To solve the problem of data shortage, we plan to develop a new type of EEG device. The new device will be a dry electrode-based wireless device similar to those manufactured by companies such as Neuroelectrics, Emotive, Cognionics, TMSi, Emotive, or Ybrain. Based on the nodes mentioned in Section 4.5, the new equipment can be made with only fewer nodes and the highly versatile device will contribute to an easy building of data. In addition, if the drowsiness detection algorithm is added to the developed mobile device, the state of drowsiness can be constantly monitored in real life and so the damages caused by daytime sleepiness can be minimized.

ACKNOWLEDGMENT

This work was supported in part by the National Research Foundation of Korea grant funded by the Ministry of Science and ICT [2018R1A2B3001628, 2014M3C9A3063541], the Ministry of Trade, Industry & Energy [10053249], the SNUBH Research Fund [13-2016-009], and the Brain Korea 21 Plus Project in 2018.

REFERENCES

[1] Patrick M. Krueger and Elliot M. Friedman. Sleep duration in the united states: A cross-sectional population-based study. *American Journal of Epidemiology*, 169 (9):1052–1063, 2009. doi: 10.1093/aje/kwp023.

[2] Jae-Hyun Kim, Kyung Ran Kim, Kyoung Hee Cho, Ki-Bong Yoo, Jeoung A Kwon, and Eun-Cheol Park. The association between sleep duration and self-rated health in the korean general population. *Journal of Clinical Sleep Medicine*, 9(10): 1057—1064, October 2013. ISSN 1550-9389. doi: 10.5664/jcsm.3082.

[3] James K. Walsh, Catherine Coulouvrat, Goeran Hajak, Matthew D. Lakoma, Maria Petukhova, Thomas Roth, Nancy A. Sampson, et al. Nighttime insomnia symptoms and perceived health in the america insomnia survey (ais). *Sleep*, 34 (8):997–1011, 2011. doi: 10.5665/SLEEP.1150.

[4] JinKwan Kim, KwangHo In, JeHyeong Kim, SeHwa You, KyungHo Kang, JaeJeong Shim, SangYeub Lee, et al. Prevalence of sleep-disordered breathing in middle-aged korean men and women. *American Journal of Respiratory and Critical Care Medicine*, 170(10):1108—1113, November 2004. ISSN 1073-449X. doi: 10.1164/rccm.200404-519oc.

[5] Young Hwangbo, Won Joo Kim, Min Kyung Chu, Chang Ho Yun, and Kwang Ik Yang. Habitual sleep duration, unmet sleep need, and excessive daytime sleepiness in korean adults. *Journal of Clinical Neurology*, 12(2):194—200, April 2016. ISSN 1738-6586. doi: 10.3988/jcn.2016.12.2.194.

[6] Sooyeon Suh, Hae-Chung Yang, Christopher P Fairholme, Hyun Kim, Rachel Manber, and Chol Shin. Who is at risk for having persistent insomnia symptoms? a longitudinal study in the general population in korea. *Sleep Medicine*, 15(2): 180—186, February 2014. ISSN 1389-9457. doi: 10.1016/j.sleep.2013.09.024.

[7] SoonJae Joo, Inkyung Baik, Hyeryeon Yi, Kihwan Jung, Jehyeong Kim, and Chol Shin. Prevalence of excessive daytime sleepiness and associated factors in the adult population of korea. *Sleep Medicine*, 10(2):182 – 188, 2009. ISSN 1389-9457. doi: https://ddoi.org/10.1016/j.sleep.2008.03.017.

[8] Chang-Ho Yun, Hyun Kim, Seung Ku Lee, Sooyeon Suh, Seung Hoon Lee, Seong-Ho Park, Robert J Thomas, et al. Daytime sleepiness associated with poor sustained attention in middle and late adulthood. *Sleep Medicine*, 16(1):143—151, January 2015. ISSN 1389-9457. doi: 10.1016/j.sleep.2014.07.028.

[9] Shantha M. W. Rajaratnam, Mark E. Howard, and Ronald R. Grunstein. Sleep loss and circadian disruption in shift work: health burden and management. *The Medical Journal of Australia*, 199(8):S11—5, October 2013. ISSN 0025-729X. doi: 10.5694/mja13.10561.

[10] Meagan Daley, Charles M. Morin, Mélanie LeBlanc, Jean-Pierre Grégoire, and Josée Savard. The economic burden of insomnia: Direct and indirect costs for

individuals with insomnia syndrome, insomnia symptoms, and good sleepers. *Sleep*, 32(1):55–64, 2009. doi: 10.5665/sleep/32.1.55.

[11] David R. Hillman and Leon C. Lack. Public health implications of sleep loss: the community burden. *The Medical Journal of Australia*, 199(8):S7—10, October 2013. ISSN 0025-729X. doi: 10.5694/mja13.10620.

[12] Yu Jin Lee, Juhyun Park, Soohyun Kim, Seong-Jin Cho, and Seog Ju Kim. Academic performance among adolescents with behaviorally induced insufficient sleep syndrome. *Journal of Clinical Sleep Medicine*, 11(1):61—68, January 2015. ISSN 1550-9389. doi: 10.5664/jcsm.4368.

[13] Megan L. Zeek, Matthew J. Savoie, Matthew Song, Leanna M. Kennemur, Jingjing Qian, Paul W. Jungnickel, and Salisa C. Westrick. Sleep duration and academic performance among student pharmacists. *American Journal of Pharmaceutical Education*, 79(5):63, June 2015. ISSN 0002-9459. doi: 10.5688/ajpe79563.

[14] So Young Kim, Sung-Gyun Kim, Songyong Sim, Bumjung Park, and Hyo Geun 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*, 95 (4):e2397, January 2016. ISSN 0025-7974. doi: 10.1097/md.0000000000002397.

[15] Jennie Connor, Robyn Norton, Shanthi Ameratunga, Elizabeth Robinson, Ian Civil, Roger Dunn, John Bailey, and Rod Jackson. Driver sleepiness and risk of serious injury to car occupants: population based case control study. *British Medical Journal (Clinical research ed.)*, 324(7346):1125, May 2002. ISSN 0959-8138. doi: 10.1136/bmj.324.7346.1125.

[16] Takashi Abe, Daniel Mollicone, Mathias Basner, and David F. Dinges. Sleepiness and safety: Where biology needs technology. *Sleep and Biological Rhythms*, 12(2): 74–84, 2014. doi: 10.1111/sbr.12067.

[17] Pia M. Forsman, Bryan J. Vila, Robert A. Short, Christopher G. Mott, and Hans P.A. Van Dongen. Efficient driver drowsiness detection at moderate levels of drowsiness. *Accident Analysis and Prevention*, 50:341 – 350, 2013. ISSN 0001-4575. doi: https://ddoi.org/10.1016/j.aap.2012.05.005.

[18] David Sandberg, Torbjörn Akerstedt, Anna Anund, Göran Kecklund, and Mattias Wahde. Detecting driver sleepiness using optimized nonlinear combinations of sleepiness indicators. *IEEE Transactions on Intelligent Transportation Systems*, 12 (1):97–108, March 2011. ISSN 1524-9050. doi: 10.1109/TITS.2010.2077281.

[19] Tang-Hsien Chang, Chih-Sheng Hsu, Chieh Wang, and Li-Kai Yang. Onboard measurement and warning module for irregular vehicle behavior. *IEEE Transactions on Intelligent Transportation Systems*, 9(3):501–513, Sept 2008. ISSN 1524-9050. doi: 10.1109/TITS.2008.928243.

[20] Suzanne Ftouni, Tracey L. Sletten, Mark Howard, Clare Anderson, Michael G. Lenné, Steven W. Lockley, and Shantha M. W. Rajaratnam. Objective and subjective measures of sleepiness, and their associations with on-road driving events in shift workers. *Journal of Sleep Research*, 22(1):58—69, February 2013. ISSN 0962-1105. doi: 10.1111/j.1365-2869.2012.01038.x.

[21] David F. Dinges and John W. Powell. Microcomputer analyses of performance on a portable, simple visual rt task during sustained operations. *Behavior Research Methods, Instruments, & Computers*, 17(6):652–655, Nov 1985. ISSN 1532-5970. doi: 10.3758/BF03200977.

[22] David F. Dinges, Naomi L. Rogers, Jillian Dorrian, and Clete A. Kushida. *Psychomotor vigilance performance: Neurocognitive assay sensitive to sleep loss.*, pages 39–70. Marcel Dekker, Inc.,, New York, NY, 2005.

[23] Julian Lim and David F. Dinges. Sleep deprivation and vigilant attention. *Annals of the New York Academy of Sciences*, 1129:305—322, 2008. ISSN 0077-8923. doi: 10.1196/annals.1417.002.

[24] Mathias Basner, Daniel Mollicone, and David F. Dinges. Validity and sensitivity of a brief psychomotor vigilance test (pvt-b) to total and partial sleep deprivation. *Acta Astronautica*, 69(11):949 – 959, 2011. ISSN 0094-5765. doi: https://doi.org/10.1016/j.actaastro.2011.07.015.

[25] David F. Dinges, Malissa M. Mallis, Greg Maislin, and John W. Powell. *Evaluation of techniques for ocular measurement as an index of fatigue and as the basis for alertness management.* Report no. DOT HS 808 762. 1998.

[26] Melinda L. Jackson, Gerard A. Kennedy, Catherine Clarke, Melissa Gullo, Philip Swann, Luke A. Downey, Amie C. Hayley, et al. The utility of automated measures of ocular metrics for detecting driver drowsiness during extended wakefulness. *Accident Analysis and Prevention*, 87:127 – 133, 2016. ISSN 0001-4575. doi: https://ddoi.org/10.1016/j.aap.2015.11.033.

[27] Eugene Aidman, Carolyn Chadunow, Kayla Johnson, and John Reece. Real-time driver drowsiness feedback improves driver alertness and self-reported driving performance. *Accident Analysis and Prevention*, 81:8 – 13, 2015. ISSN 0001-4575. doi: https://ddoi.org/10.1016/j.aap.2015.03.041.

[28] Mark Corbett. A drowsiness detection system for pilots: Optalert. *Aviation, Space, and Environmental Medicine*, 80(2):149, February 2009. ISSN 0095-6562. doi: 10.3357/asem.21001.2009.

[29] Margherita Fabbri, Federica Provini, Elisa Magosso, Anna Zaniboni, Antonietta Bisulli, Giuseppe Plazzi, Mauro Ursino, and Pasquale Montagna. Detection of sleep onset by analysis of slow eye movements: A preliminary study of mslt recordings. *Sleep Medicine*, 10(6):637 – 640, 2009. ISSN 1389-9457. doi: https://ddoi.org/10.1016/j.sleep.2008.07.007.

[30] Jim A. Horne and Louise A. Reyner. Counteracting driver sleepiness: effects of napping, caffeine, and placebo. *Psychophysiology*, 33(3):306–309, 1996. ISSN 0048-5772.

[31] Sarah McGuire, Uwe Müller, Eva-Maria Elmenhorst, and Mathias Basner. Inter-individual differences in the effects of aircraft noise on sleep fragmentation. *Sleep*, 39(5):1107–1110, 2016. doi: 10.5665/sleep.5764.

[32] Linda Ng Boyle, Jon Tippin, Amit Paul, and Matthew Rizzo. Driver performance in the moments surrounding a microsleep. *Transportation Research Part F: Traffic Psychology and Behaviour*, 11(2):126 – 136, 2008. ISSN 1369-8478. doi: https://ddoi.org/10.1016/j.trf.2007.08.001.

[33] J. Faber. Detection of different levels of vigilance by eeg pseudo spectra. *Neural Network World*, 14:285–290, 01 2004. ISSN 1210-0552.

[34] Agustina Garcés Correa, Lorena Orosco, and Eric Laciar. Automatic detection of drowsiness in eeg records based on multimodal analysis. *Medical Engineering and Physics*, 36(2):244–249. ISSN 1350-4533. doi: 10.1016/j.medengphy.2013.07.011.

[35] Hyungseob Han and Kyoung-Young Song. Electroencephalogram-based driver drowsiness detection system using errors-in-variables (eiv) and multilayer perceptron (mlp). *The Journal of Korean Institute of Communications and Information Sciences*, 39(10):887–895, 2014.

[36] Tzyy-Ping Jung, S. Makeig, M. Stensmo, and T. J. Sejnowski. Estimating alertness from the eeg power spectrum. *IEEE Transactions on Biomedical Engineering*, 44 (1):60–69, Jan 1997. ISSN 0018-9294. doi: 10.1109/10.553713.

[37] Zahra Mardi, Seyedeh Naghmeh Miri Ashtiani, and Mohammad Mikaili. Eeg-based drowsiness detection for safe driving using chaotic features and statistical tests. *Journal of Medical Signals and Sensors*, 1(2):130–137, 2011. ISSN 2228-7477.

[38] Nikhil R. Pal, Chien-Yao Chuang, Li-Wei Ko, Chih-Feng Chao, Tzyy-Ping Jung, Sheng-Fu Liang, and Chin-Teng Lin. Eeg-based subject- and session-independent drowsiness detection: An unsupervised approach. *EURASIP Journal on Advances in Signal Processing*, 2008:192:1–192:11, January 2008. ISSN 1110-8657.

[39] Christos Papadelis, Zhe Chen, Chrysoula Kourtidou-Papadeli, Panagiotis D. Bamidis, Ioanna Chouvarda, Evangelos Bekiaris, and Nikos Maglaveras. Monitoring sleepiness with on-board electrophysiological recordings for preventing sleep-deprived traffic accidents. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 118(9):1906—1922, September 2007. ISSN 1388-2457. doi: 10.1016/j.clinph.2007.04.031.

[40] Arun Sahayadhas, Kenneth Sundaraj, and Murugappan Murugappan. Drowsiness detection during different times of day using multiple features. *Australasian Physical & Engineering Sciences in Medicine*, 36(2):243–250, Jun 2013. ISSN 1879-5447. doi: 10.1007/s13246-013-0200-6.

[41] Abdulhamit Subasi. Automatic recognition of alertness level from eeg by using neural network and wavelet coefficients. *Expert Systems with Applications*, 28(4): 701 – 711, 2005. ISSN 0957-4174.

[42] Mervyn V.M. Yeo, Xiaoping Li, Kaiquan Shen, and Einar P.V. Wilder-Smith. Can svm be used for automatic eeg detection of drowsiness during car driving? *Safety Science*, 47(1):115 – 124, 2009. ISSN 0925-7535. doi: https://ddoi.org/10.1016/j.ssci.2008.01.007.

[43] H. S. Choi, B. Lee, and S. Yoon. Biometric authentication using noisy electrocardiograms acquired by mobile sensors. *IEEE Access*, 4:1266–1273, 2016. doi: 10.1109/ACCESS.2016.2548519.

[44] S. Choi, S. Kim, J. Seo, J. Y. Park, and S. Yoon. Wearable and wireless measurement system for evaluating penile tumescence. In *2015 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, pages 1–4, Oct 2015. doi: 10.1109/BioCAS.2015.7348287.

[45] Maxim Y. Khitrov, Srinivas Laxminarayan, David Thorsley, Sridhar Ramakrishnan, Srinivasan Rajaraman, Nancy J. Wesensten, and Jaques Reifman. Pc-pvt: A platform for psychomotor vigilance task testing, analysis, and prediction. *Behavior Research Methods*, 46(1):140–147, Mar 2014. ISSN 1554-3528. doi: 10.3758/s13428-013-0339-9.

[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 *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 5820–5823, Aug 2014. doi: 10.1109/EMBC.2014.6944951.

[47] Arcady A. Putilov and Olga G. Donskaya. Construction and validation of the eeg analogues of the karolinska sleepiness scale based on the karolinska drowsiness test. *Clinical Neurophysiology*, 124(7):1346 – 1352, 2013. ISSN 1388-2457. doi: https://ddoi.org/10.1016/j.clinph.2013.01.018.

[48] Christi E. D. Alloway, Robert D. Ogilvie, and Colin M. Shapiro. The alpha attenuation test: Assessing excessive daytime sleepiness in narcolepsy-cataplexy. *Sleep*, 20(4):258–266, 1997. doi: 10.1093/sleep/20.4.258.

[49] Emma M. Whitham, Kenneth J. Pope, Sean P. Fitzgibbon, Trent Lewis, C. Richard Clark, Stephen Loveless, Marita Broberg, et al. Scalp electrical recording during paralysis: Quantitative evidence that eeg frequencies above 20hz are contaminated by emg. *Clinical Neurophysiology*, 118(8):1877 – 1888, 2007. ISSN 1388-2457. doi: https://ddoi.org/10.1016/j.clinph.2007.04.027.

[50] Chin-Teng Lin, Ruei-Cheng Wu, Sheng-Fu Liang, Wen-Hung Chao, Yu-Jie Chen, and Tzyy-Ping Jung. Eeg-based drowsiness estimation for safety driving using independent component analysis. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 52(12):2726–2738, Dec 2005. ISSN 1549-8328. doi: 10.1109/TCSI.2005.857555.

[51] Sheng-Hsiou Hsu and Tzyy-Ping Jung. Monitoring alert and drowsy states by modeling eeg source nonstationarity. *Journal of Neural Engineering*, 14(5):056012, 2017.

[52] Pai-Yuan Tsai, Weichih Hu, Terry BJ Kuo, and Liang-Yu Shyu. A portable device for real time drowsiness detection using novel active dry electrode system. In *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 3775–3778, Sept 2009. doi: 10.1109/IEMBS.2009.5334491.

[53] Gang Li and Wan-Young Chung. Estimation of eye closure degree using eeg sensors and its application in driver drowsiness detection. *Sensors*, 14(9):17491–17515, 2014. ISSN 1424-8220. doi: 10.3390/s140917491.

[54] Muhammad Awais, Nasreen Badruddin, and Micheal Drieberg. A hybrid approach to detect driver drowsiness utilizing physiological signals to improve system performance and wearability. *Sensors*, 17(9), 2017.

[55] Daniel Ribeiro, César Teixeira, and Alberto Cardoso. Eeg-based drowsiness detection platform to compare different methodologies. In *2017 4th Experiment@International Conference*, pages 318–322, June 2017. doi: 10.1109/EXPAT.2017.7984343.

[56] Peter Welch. The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on Audio and Electroacoustics*, 15(2):70–73, Jun 1967. ISSN 0018-9278. doi: 10.1109/TAU.1967.1161901.

[57] David J. Thomson. Spectrum estimation and harmonic analysis. *Proceedings of the IEEE*, 70(9):1055–1096, Sept 1982. ISSN 0018-9219. doi: 10.1109PROC.1982.12433.

[58] Michael J. Prerau, Ritchie E. Brown, Matt T. Bianchi, Jeffrey M. Ellenbogen, and Patrick L. Purdon. Sleep neurophysiological dynamics through the lens of multitaper spectral analysis. *Physiology*, 32(1):60–92, 2017. doi: 10.1152/physiol.00062.2015. PMID: 27927806.

[59] Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, pages 785–794, New York, NY, USA, 2016. ACM. ISBN 978-1-4503-4232-2. doi: 10.1145/2939672.2939785.

[60] Claire Adam-Bourdarios, G. Cowan, Cecile Germain-Renaud, Isabelle Guyon, Balázs Kégl, and D. Rousseau. The higgs machine learning challenge. *Journal of Physics*, 664(7):072015, 2015.

[61] Jerome H. Friedman. Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5):1189–1232, 2001. ISSN 00905364.

[62] Elizabeth R. DeLong, David M. DeLong, and Daniel L. Clarke-Pearson. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*, 44(3):837–845, 1988. doi: 10.2307/2531595. cited By 7278.

[63] Wu Ting, Yan Guo-zheng, Yang Bang-hua, and Sun Hong. Eeg feature extraction based on wavelet packet decomposition for brain computer interface. *Measurement*, 41(6):618 – 625, 2008. ISSN 0263-2241.

[64] Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, Oct 2001. ISSN 1573-0565. doi: 10.1023/A:1010933404324.

[65] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, Sep 1995. ISSN 1573-0565. doi: 10.1007/BF00994018.

[66] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521:436, 2015. doi: 10.1038/nature14539.

[67] Nitesh V. Chawla, Kevin W. Bowyer, Lawrence O. Hall, and W. Philip Kegelmeyer. Smote: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16:321–357, 2002.

[68] Seonwoo Min, Byunghan Lee, and Sungroh Yoon. Deep learning in bioinformatics. *Briefings in Bioinformatics*, 18(5):851–869, 2017. doi: 10.1093/bib/bbw068.

[69] Junghwan Baek, Byunghan Lee, Sunyoung Kwon, and Sungroh Yoon. Lncrnanet: long non-coding rna identification using deep learning. *Bioinformatics*, page bty418, 2018. doi: 10.1093/bioinformatics/bty418.

[70] Taesup Moon, Seonwoo Min, Byunghan Lee, and Sungroh Yoon. Neural universal discrete denoiser. In D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, editors, *Advances in Neural Information Processing Systems 29*, pages 4772–4780. Curran Associates, Inc., 2016.

[71] Sunyoung Kwon and Sungroh Yoon. Deepcci: End-to-end deep learning for chemical-chemical interaction prediction. In *Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology,and Health Informatics*, ACM-BCB '17, pages 203–212, New York, NY, USA, 2017. ACM. ISBN 978-1-4503-4722-8. doi: 10.1145/3107411.3107451.

[72] Seunghyun Park, Seonwoo Min, Hyun-Soo Choi, and Sungroh Yoon. Deep recurrent neural network-based identification of precursor microrna. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems 30*, pages 2891–2900. Curran Associates, Inc., 2017.

[73] Byunghan Lee, Junghwan Baek, Seunghyun Park, and Sungroh Yoon. deeptarget: End-to-end learning framework for microrna target prediction using deep recurrent neural networks. In *Proceedings of the 7th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*, BCB '16, pages 434–442, New York, NY, USA, 2016. ACM. ISBN 978-1-4503-4225-4. doi: 10.1145/2975167.2975212.

[74] Hui Kwon Kim, Seonwoo Min, Myungjae Song, Soobin Jung, Jae Woo Choi, Younggwang Kim, Sangeun Lee, et al. Deep learning improves prediction of crispr–cpf1 guide rna activity. *Nature biotechnology*, 36(3):239, 2018. doi: http://dx.doi.org/10.1038/nbt.4061.