Neurocomputing 408 (2020) 100–111

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Driver sleepiness detection from EEG and EOG signals using GAN and LSTM networks

Yingying Jiao[a,1], Yini Deng[a,1], Yun Luo[a], Bao-Liang Lu[a,b,c,*]

[a] Center for Brain-like Computing and Machine Intelligence, Department of Computer Science and Engineering, Shanghai Jiao Tong University, 800 Dong Chuan Road, Shanghai 200240, China
[b] Key Laboratory of Shanghai Education Commission for Intelligent Interaction and Cognition Engineering, Shanghai Jiao Tong University, 800 Dong Chuan Road, Shanghai 200240, China
[c] Brain Science and Technology Research Center, Shanghai Jiao Tong University, 800 Dong Chuan Road, Shanghai 200240, China

ARTICLE INFO

Article history:
Received 9 February 2019
Revised 6 May 2019
Accepted 24 May 2019
Available online 10 March 2020

Keywords:
EEG
EOG
Alpha blocking
Alpha wave attenuation-disappearance
GAN
LSTM
Sleepiness detection

ABSTRACT

In recent years, sleepiness during driving has become a main cause for traffic accidents. However, the fact is that we know very little yet about the electrophysiological marker for assessing diver sleepiness. Previous studies and our researches have shown that alpha blocking phenomenon and alpha wave attenuation-disappearance phenomenon represent two different sleepiness levels, the relaxed wakefulness and the sleep onset, respectively. This paper proposes a novel model for driver sleepiness detection based on electroencephalography (EEG) and electrooculography (EOG) signals. Our model aims to track the change in alpha waves and differentiate the two alpha-related phenomena. Continuous wavelet transform is adopted to extract features from physiological signals in both time and frequency domains. Meanwhile, Long-Short Term Memory (LSTM) network is introduced to deal with temporal information of EEG and EOG signals. To deal with insufficient physiological sample problem, generative adversarial network (GAN) is used to augment the training dataset. Experimental results indicate that the F1 score for detecting start and end points of alpha waves reaches to around 95%. And Conditional Wasserstein GAN (CWGAN) we adopted was effective in augmenting dataset and boost classifier performance. Meanwhile, our LSTM classifier achieved a mean accuracy of 98% for classifying end points of alpha waves under leave-one-subject-out cross validation.

1. Introduction

In recent years, driver sleepiness has become one of the main reasons for traffic accidents. Many drivers have the experience of feeling drowsy or falling asleep while driving the car, but most of them tend to ignore the early signs of sleepiness [1–3]. Especially for train drivers and pilots, this may cause more casualty. Since sleepiness level transition is a physiological process, physiological signals change obviously with it [4–7]. Therefore, finding a reliable electrophysiological marker and a detection model for it is of practical importance.

There are mainly four different ways to detecting driver sleepiness: (1) self-evaluation; (2) vehicle-based detection; (3) behavioral measures; and (4) physiological signals [7–10]. And to have a greater potential for routine use in real-life scenario, a driver sleepiness detection system should be practical, objective and robust. However, driver's self evaluation is obviously subjective. For vehicle-based detection, since it often makes use of steering wheel movements and other input from sensors, it is difficult to develop a general model due to complex geometric characteristic of roads and various vehicle types [4,11]. And behavioral measures, such as PERCLOS (the percentage of eye closure), are often limited to driving conditions like illumination. Compared to these measures, physiological signals can reflect the mechanism of driver sleepiness in nature [5–7]. Meanwhile, EEG is regarded as the 'gold standard' and widely used as an indicator of the transition between wakefulness and sleep [4–8].

Spontaneous alpha activity observed on EEG signals can reflect different underlying physiological process [5,7,8]. Alpha waves originate from the occipital lobe, and can be observed under the relaxed wakefulness state during eye closure. This refers to the typical alpha blocking phenomenon: when we close our eyes, alpha waves continuously appear; when we reopen our

* Corresponding author at: Center for Brain-like Computing and Machine Intelligence, Department of Computer Science and Engineering, Shanghai Jiao Tong University, 800 Dong Chuan Road, Shanghai 200240, China.
E-mail address: bllu@sjtu.edu.cn (B.-L. Lu).
[1] The first two authors contributed equally to this work and should be considered as joint first authors.

https://doi.org/10.1016/j.neucom.2019.05.108
0925-2312/

eyes, alpha waves quickly disappear [7,12–14]. Meanwhile, in our previous research, another alpha-related phenomenon, alpha wave attenuation-disappearance phenomenon, has been found and proven to be practical for predicting the sleep onset state [15,16]. Under higher sleepiness state, when we close our eyes, alpha waves attenuate until they finally disappear. The attenuation of alpha waves is the most valid electrophysiological marker for the beginning of sleep onset period (SOP), which refers to the transition from wakefulness to real sleep [5,17]. This transition is precisely described in Hori's nine-stage system, which has been validated to be particularly useful for describing SOP. Compared to a train of alpha waves in Hori's sleep stage 1 (wakefulness), alpha wave activity decrease gradually from more than 50% in sleep stage 2 to less than 50% in sleep stage 3 [5].

In our simulated driving experiments, driver sleepiness is caused by long-time continuous driving task and partial sleep deprivation. The typical alpha blocking phenomenon and alpha wave attenuation-disappearance phenomenon occur during eye-closed period, representing two different sleepiness levels, the relaxed wakefulness and the sleep onset, respectively. By differentiating these two phenomena, when a subject closes his/her eyes, we can determine the two sleepiness levels. Therefore, our study first tracks the change in alpha waves to detect end points of alpha waves in the two alpha related phenomena, and then the corresponding EOG signals are used to distinguish the two phenomena.

In this paper, we propose a novel method that combines continuous wavelet transform (CWT) with Long-Short Term Memory (LSTM) to detect drive sleepiness from EEG and EOG signals. When the two alpha-related phenomena occur, EEG and EOG signals both change correspondingly. Thus, we utilize CWT to represent the change of EEG and EOG signals in time and frequency domains. At the same time, LSTM is introduced to classify different waveform in EOG signals which have inherent temporal dependencies. In our previous work, we used SVM for classification, but it does not take temporal dependencies into account. Besides the application in the field of computer vision [18–21], LSTM, as a special form of recurrent neural network, is an appropriate deep learning architecture for analyzing biomedical signals [22–25]. Meanwhile, it is capable of handling long-term dependencies and can avoid the vanishing gradient problem. However, the training of LSTM requires large number of samples, whereas it is usually difficult to get sufficient samples from physiological signals. To solve this problem, Conditional Wasserstein GAN (CWGAN) is adopted to augment EOG data for training the LSTM classifier [26]. By adversarial training, GANs have strong abilities to generate realistic-like data, and are widely applied to images [27].

The main contributions of this paper to driver sleepiness detection from EEG and EOG signals can be summarized as follows:

(1) We utilize CWT with proper basis functions to characterize the peculiarities in EEG and EOG signals.
(2) We make use of CWGAN for EOG data augmentation to train the LSTM classifier.
(3) LSTM is adopted to learn the temporal information in EOG signals in order to classify different waveform.
(4) Our proposed method places very few electrodes on subjects, and it is thus a promising approach to implementing driver sleepiness detection in real-life scenarios.

2. Related work

2.1. Detecting driver sleepiness using physiological signals

Among all physiological signals, EEG is regarded as the most reliable signal for detecting sleepiness. Numerous studies have demonstrated that EEG signals such as alpha (8–12 Hz), theta (4–7 Hz) and delta (1–3 Hz) are highly correlated with driver sleepiness [28,29]. Li et al. developed an EEG-based driver drowsiness detection system with an SVM-based posterior probabilistic model and a wearable EEG headband [30]. Shi and Lu proposed an EEG-based vigilance estimation model using extreme learning machines [31]. Wu et al. proposed a novel online weighted adaptation regularization for regression algorithm to estimate driver drowsiness from EEG signals [32]. Lin et al. used EEG spectra to predict driver drowsiness [33].

In addition to EEG, EOG, which contains characteristic information of eye blinks and other eye movements, is also widely used to estimate sleepiness due to its high signal-noise ratio. Jammes et al. proposed an automatic detection algorithm for eye blinks in EOG to estimate drowsiness according to blink duration [34]. Ma et al. used EOG features, mainly slow eye movements, to estimate the human vigilance changes [35]. Zhang et al. proposed a novel approach to driving fatigue detection using forehead EOG, and SVM's prediction correlation coefficient reached 0.88 on average [36].

Meanwhile, many researches combined EEG and EOG signals to increase the reliability and accuracy of sleepiness detection. Zheng and Lu combined EEG and forehead EOG to leverage their complementary characteristics for vigilance estimation [37]. Kartsch et al. developed a wearable EEG-based drowsiness detection system by determining blink durations and recognizing the presence of alpha waves [38]. Using both EEG and EOG signals, Chen et al. combine the wavelet-based nonlinear features with extreme learning machine classifier for detecting drowsy/alert states [39]. Arnin et al. developed a portable wireless drowsiness detection system according to the characteristics of eye blinks and eye movements and the changes in EEG waves in drowsy state [40].

In recent years, along with the quick development of wearable devices for processing EEG and other physiological signals, the demand for methodologies based on fewer channels has been raised. Zheng et al. implemented continuous vigilance estimation using a wearable EOG device [41]. So far, fusing multiple physiological signals with fewer channels has been the main trend for designing reliable and feasible driver sleepiness detection system.

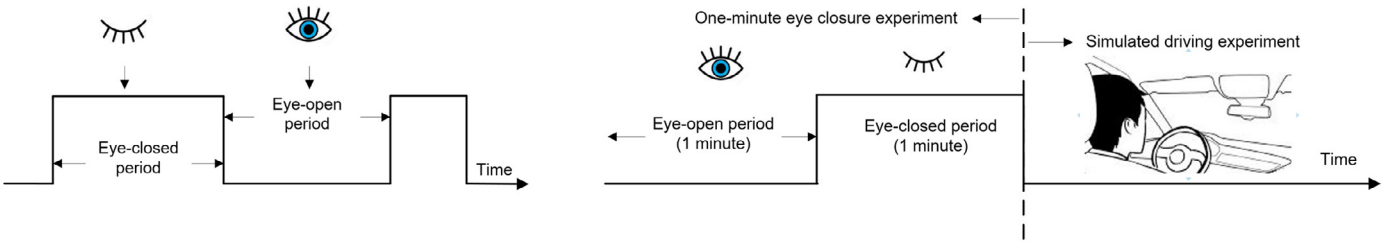
2.2. Using deep networks to classify EEG and EOG signals

Some recent works have proposed novel methods that utilize LSTM to classify EEG signals. Dong et al. applied a Mixed Neural Network approach to classify temporal sleep stage using features extracted from EEG signals [24]. Their MNN model combined a rectifier neural network with LSTM, and their method was proven to be effective compared with other existing models. Similarly, Supratak et al. proposed DeepSleepNet for automatic sleep stage scoring based on raw single-channel EEG [25]. The model mainly adopted two CNNs for feature representation and a bidirectional-LSTM for sequence residual learning, and the sleep stage scoring performance of DeepSleepNet achieved similar overall accuracy compared to the state-of-the-art hand-engineering methods. Besides, Tang et al. utilized bimodal LSTM for emotion recognition using EEG and eye movement signals, and the model outperformed other methods with a mean accuracy of 93.97% [42].

3. Materials

3.1. Experimental procedure

In total, twelve healthy subjects (8 males and 4 females, with an average age of 22) who have siesta habit for more than a year participated in our experiments. The study was approved by the local ethics committee, and subjects gave written informed consent before participation. The Epworth Sleepiness Scale (ESS) [43], a simple and reliable method for measuring the general level of



(a) Repeated eye closure experiment (b) One-minute eye closure experiment and simulated driving experiment

Fig. 1. Illustration of three kinds of experiments performed in this paper.

daytime sleepiness for adults, was used to measure the daytime sleepiness level of subjects. The mean ± SD of ESS score for all twelve subjects is 10.5 ± 1.2. For each subject, we conducted three experiments: (1) Repeated eye closure experiment; (2) One-minute eye closure experiment; and (3) Simulated driving experiment.

3.1.1. Repeated eye closure experiment

The repeated eye closure experiment aims to obtain EEG and EOG signals with alpha blocking phenomenon. As shown in Fig. 1(a), in this experiment, subjects closed and opened their eyes according to the instruction. The duration of eye closure was random, ranging from 5 s to 15 s. And the subjects were seated in a chair under relaxed wakefulness. Each subject participated in this experiment for about 60 minutes once.

3.1.2. One-minute eye closure experiment

The one-minute eye closure experiment was conducted just before the simulated driving experiment, as shown in Fig. 1(b). The subjects kept their eyes opened and closed for one minute respectively under the relaxed wakefulness state. During the one-minute eye closure, alpha waves were observed to be prominent and continuous, with a high amplitude. In contrary, when eyes were opened, we could barely observe alpha waves. Then, alpha wavelet energy threshold was calculated based on EEG signals obtained from this experiment, and the threshold was later used to track the change of alpha waves. This experiment and the simulated driving experiment were conducted in succession with the same experimental setup, so that we could maintain data consistency for subsequent data analysis and processing.

3.1.3. Simulated driving experiment

The simulated driving experiment was conducted to induce the sleepiness level change of the subject from relaxed wakefulness to sleep onset. The experiment started at the subject's usual nap time, which leads to a certain degree of sleep deprivation and thus makes it easier to induce a higher sleepiness level. Fig. 2 is our experimental environment, in which there was a real vehicle without engine. The subject was seated in the car, and operated the steering wheel and gas pedal to drive a virtual car on a highway scene shown on the LCD screen. To effectively make the subject drowsy, the virtual car's speed was slow, and the scene around the road was monotonous.

The experimental task for subjects was to drive the car properly in the virtual scene. Meanwhile, subjects were asked not to deliberately close their eyes when they were not sleepy. However, lacking real threat to life safety, subjects tended to close their eyes spontaneously. Therefore, a lot of eye closure periods appeared, during which we could observe the two alpha-related phenomena. Due to unfamiliarity with the experimental settings, a higher sleepiness level usually could not be induced when the subject participated in this experiment for the first time. Therefore, they often took part in this experiment for 2 to 3 times, and each lasted for about 90 minutes.

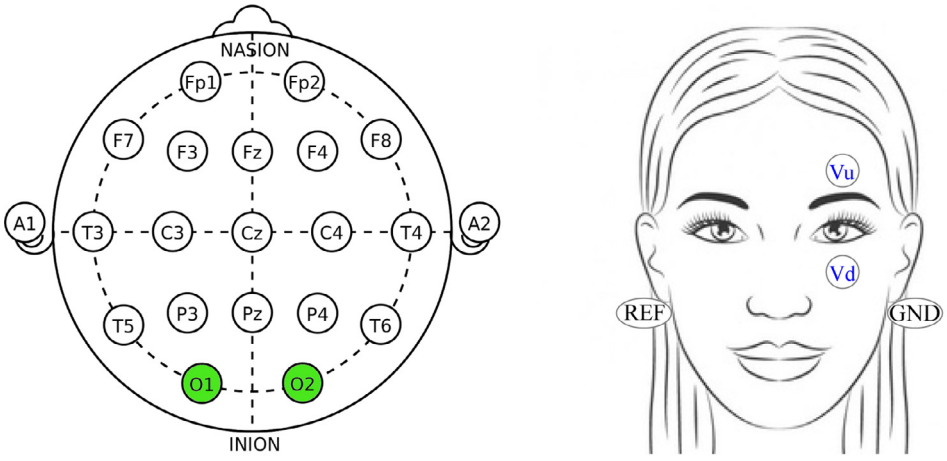
(a) Simulated driving environment

(b) A subject is participating in simulated driving experiment

Fig. 2. Environment for simulated driving experiment.

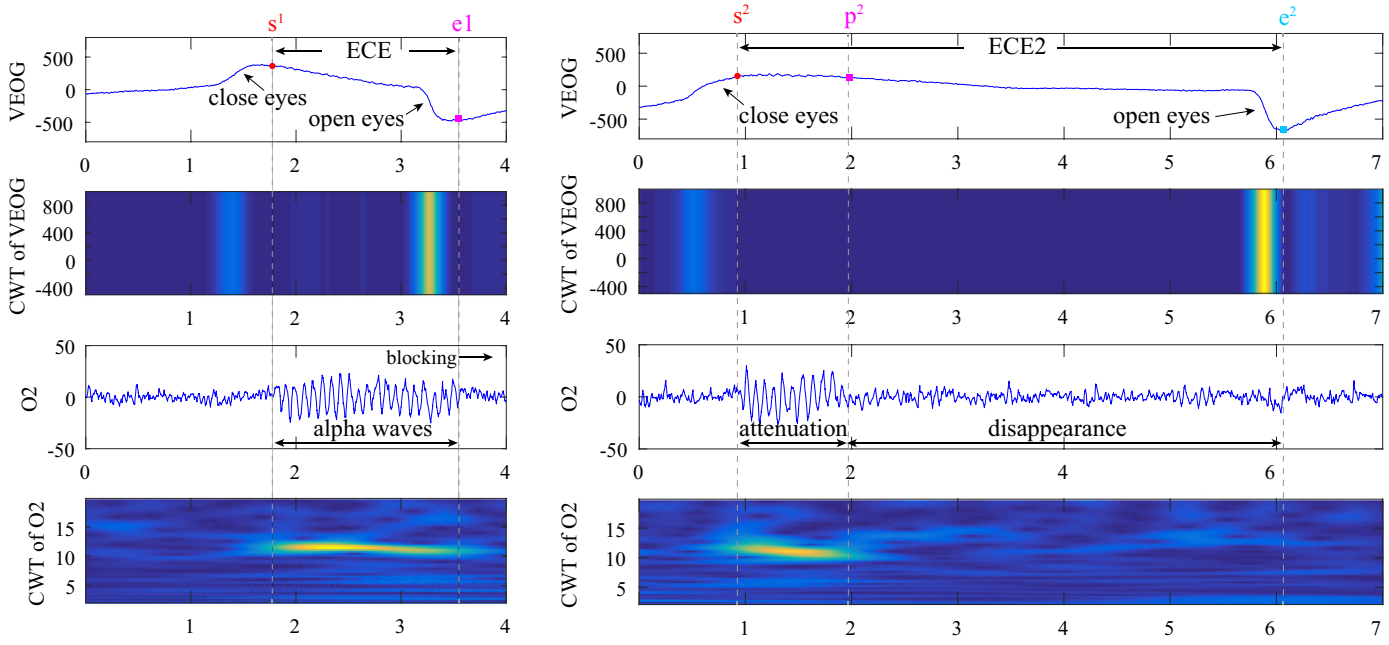
3.2. Data recording

For all the experiments, data recording settings were the same. EEG and EOG signals were recorded at a 1000 Hz sampling rate using ESI NeuroScan System. In total, six electrodes were placed on the subject, as shown in Fig. 3: (1) two electrodes, Vu and Vd, for obtaining VEOG signal, and VEOG = Vu - Vd; (2) two occipital



(a) EEG electrode placement (b) EOG electrodes placement

Fig. 3. Illustration of EEG and EOG electrode placement.



(a) Alpha blocking phenomenon (b) Alpha wave attenuation-disappearance phenomenon

Fig. 4. VEOG signal, O2 signal and their corresponding continuous wavelet transform results for the two alpha-related phenomena.

electrodes for obtaining EEG signal, and we used O1 as an alternative when O2 was corrupted by noise; and (3) one reference (REF) electrode and one ground (GND) electrode behind the ears. Subjects were asked to wear a headband to fix O1 and O2 electrodes.

Meanwhile, a digital video camera was set to monitor the subject's face. In order to observe eye movements and the corresponding physiological signals at the same time, the facial video captured from the camera and the signals displayed on the Scan software interface were synchronously recorded into a file. This also allows us to review and investigate the relationship between eye movements and EEG/EOG signals manually.

3.3. Two alpha-related phenomena and their visual marking

3.3.1. Eye closure event (ECE)

As shown in the VEOG signal in Fig. 4, the behavior of closing and opening eyes produces an upward trend line and a downward trend line, respectively. And the energy value of VEOG signal around these two trend lines is larger (vertical bars in yellow and light blue on CWT of VEOG). Based on this pattern, an eye closure event is defined as the period from the end point of the upward trend line to that of the downward trend line on VEOG signal.

3.3.2. Presence of alpha waves

Alpha waves can be observed on O2 signal. And we judged the presence of alpha waves by visual inspection. That is, when we found a sinusoidal rhythm of 8 to 12 cps, which corresponds to the frequency of alpha frequency band, we considered alpha waves to be present [44,45]. In addition, alpha rhythm was double-checked by apply CWT, because alpha wavelet energy was higher after the transform, shown as the horizontal areas in yellow on CWT of O2 in Fig. 4.

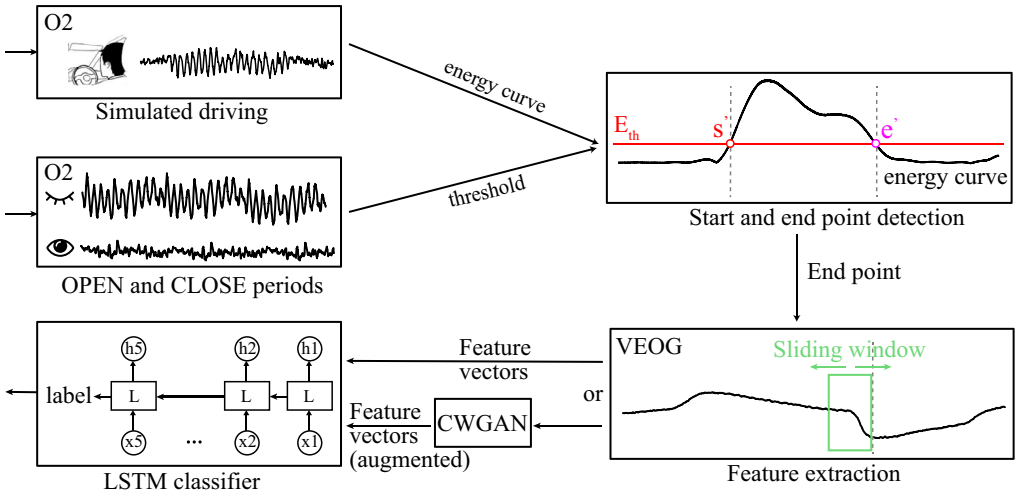


Fig. 5. The flowchart of the proposed model.

3.3.3. Alpha blocking phenomenon

Alpha blocking phenomenon is illustrated in Fig. 4(a) where spontaneous alpha oscillations are prominent and continuous on O2 signal during eye closure period. Alpha waves appear when eyes are closed under relaxed wakefulness, and quickly disappear when eyes are reopened [7]. When applying CWT to O2 signal, we can see that the CWT coefficients are large in alpha frequency band (8 - 12 Hz) within the period of alpha blocking phenomenon.

On O2 signal in Fig. 4(a), according to the way we judged the presence of alpha waves in Section 3.3.2, we visually marked s^1 and e^1 as the start and end points of alpha waves in alpha blocking phenomenon. s^1 and e^1 also correspond approximately to the end point of the upward trend line and the downward trend line in VEOG signal. Therefore, we define ECE^1 as the eye closure event corresponding to alpha blocking phenomenon.

3.3.4. Alpha wave attenuation-disappearance phenomenon

In our previous work [15,16,46], we found alpha wave attenuation-disappearance phenomenon and demonstrated that attenuation-disappearance phenomenon can be used as an indicator for the entry into sleep. As shown on O2 signal in Fig. 4(b), this phenomenon has a split point, before which alpha waves present and gradually attenuate, and after which alpha waves almost completely disappear. We consider this split point as the end point of alpha waves, and it was marked as p^2 according to the way we judged the presence of alpha waves in Section 3.3.2. Similar to alpha blocking phenomenon, the start point of alpha waves is marked as s^2, corresponding to the end point of the upward trend line on VEOG signal. In addition to s^2, we marked e^2 as the end point of the downward trend line on VEOG signal based on the pattern of ECE in Section 3.3.1. Therefore, with s^2 and e^2, we define ECE^2 as the eye closure event corresponding to alpha wave attenuation-disappearance phenomenon.

4. Methods

As the two alpha-related phenomena represent two different sleepiness levels, the key to detecting driver sleepiness is to distinguish these two phenomena. Therefore, our proposed model tracks the change in alpha waves on O2 signal, and utilizes VEOG signal for classification.

Fig. 5 is the flowchart of our model, and Algorithm 1 describes the main steps of constructing the proposed model. First, we applied CWT on O2 signal from the one-minute eye closure experiment in order to obtain alpha wavelet energy threshold E_{th}. Similarly, we applied CWT on O2 signal from the simulated driving experiment so that we could get a wavelet energy curve over time. Then, alpha wavelet energy threshold was compared with the curve so as to detect the start and end points of alpha waves. If the detected point was an end point, we first extracted features from its corresponding VEOG signal using CWT. Then, depending on different training and test strategies, the features may be augmented using CWGAN framework. Finally, features were put into the trained LSTM classifier to determine whether the corresponding end point is the end point of alpha waves in alpha blocking phenomenon or in alpha wave attenuation-disappearance phenomenon.

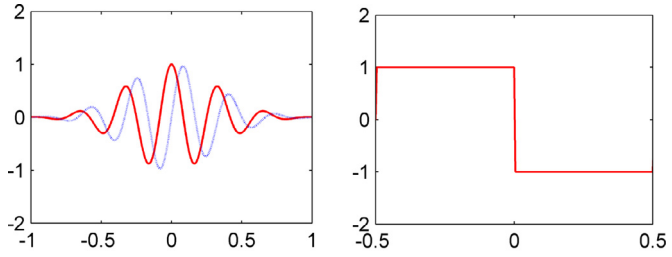
Algorithm 1 Driver Sleepiness Detection.

```
E_max: maximum alpha wavelet energy value during OPEN period
E_min: minimum alpha wavelet energy value during CLOSE period
Calculate alpha wavelet energy threshold E_th = (E_max + E_min)/2
Calculate alpha wavelet energy E_w during simulated driving experiment
Train the LSTM classifier using (augmented and) labeled VEOG features
for all time point t in E_w do
    Compare E_w at time point t with E_th
    if time point t_e is an end point of alpha waves then
        Extract features from VEOG signals around t_e
        Put the extracted features into the trained LSTM classifier
        if the extracted features belong to the end point of alpha waves in alpha wave attenuation-disappearance phenomenon then
            The subject is under the sleep onset state
        else
            The subject is under the relaxed wakefulness state
        end if
    end if
end for
```

4.1. Obtaining alpha wavelet energy threshold

4.1.1. Continuous wavelet transform (CWT)

As a signal processing tool, continuous wavelet transform has the following advantages for processing EEG and EOG signals:



(a) Complex Morlet wavelet (b) Haar wavelet

Fig. 6. Mother wavelets for CWT. (a) The red and blue curve represent the real and imaginary components, respectively. (b) Haar Wavelet.

(1) It reveals the time-frequency structure of signals, thus allowing localization of their peculiar features in time and frequency domains.
(2) It provides more flexibility in the choice of basis function into which a signal is expanded.
(3) It enables efficacious analysis of short-term time series containing the abrupt transitions or characteristic oscillation periods.

Therefore, we utilized CWT for time-frequency analysis of EEG and EOG signals to deal with their non-stationary properties [47,48]. CWT measures the similarity between an input signal $f(t)$ and the analyzing function, which is a mother wavelet, as follows,

$$W(s,\tau) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} f(t)\psi^*(\frac{t-\tau}{s})dt \quad (1)$$

where s is the time scale, τ is the time shift, ψ is the basic complex function (or mother wavelet) of the wavelet transform, and * denotes its complex conjugate.

According to Eq. (1), CWT compares the input signal to the shifted and compressed or stretched version of a mother wavelet. By varying the value of s and τ in a continuous way, we can obtain CWT coefficient $W(s, \tau)$. And wavelet coefficients represent the degree of correlations between mother wavelet and the input signal [49]. Meanwhile, the better the wavelet function matches the input signal, the more accurately the features of the input signal can be represented by wavelet coefficients [49,50].

We selected complex Morlet wavelet as the mother wavelet for O2 signals, because it has a similar geometric shape to alpha waves. Complex Morlet wavelet is a complex exponential modulated Gaussian function, as shown in Fig. 6(a).

According to the scale-frequency relationship for complex Morlet wavelet, wavelet energy $w(t)$ in the fixed alpha frequency band F = [8, 12] Hz was calculated using wavelet coefficient $W(f_s, t)$ as follows,

$$w(t) = \int_F |W(f_s,t)|^2 df_s \quad (2)$$

where $W(f_s, t)$ is equivalent to $W(s, \tau)$ in Eq. (1). Then, $w(t)$ was averaged in time window T,

$$E_w = \frac{\int_{t-T/2}^{t+T/2} w(t)dt}{F_s * T} \quad (3)$$

where F_s is the sampling rate. And T was set to 1 s, because obvious short burst of alpha waves hold for approximately 1 s.

4.2. Calculating alpha wavelet energy threshold

O2 signal in the one-minute eye closure experiment was used to calculate alpha wavelet energy threshold for each subject. Empirically, CWT with complex Morlet mother wavelet scaling from 1 to 1024 was applied to the one-minute eye-open period and one-minute eye-closed period respectively. According to Eq. (3), window length was set to 1 s and sliding step size was set to 0.1 s. So we obtained numerous alpha wavelet energy values for the two one-minute periods. Then, in order to avoid interference with noise, we removed those alpha wavelet energy values that were more than the variance of the distribution from the two periods. The distribution of alpha wavelet energy value was shown in Fig. 7. Finally, we found the minimum energy value during the eye-closed period and the maximum during eye-open period. In theory, the threshold can take any value between them. However, noise is usually included in O2 signals, so we calculated the mean of those two values as the threshold. Meanwhile, according to our observation and calculation, the threshold is different across subject.

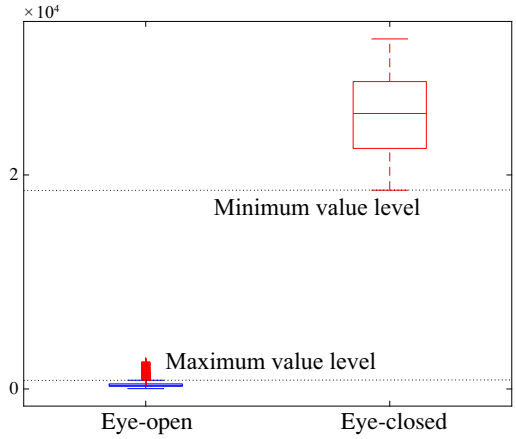


Fig. 7. Alpha wavelet energy distribution during OPEN and CLOSE period. This is one subject's data.

4.3. Detecting start and end points of alpha waves

As illustrated on E_w in Fig. 8, to obtain alpha wavelet energy value curve, we applied CWT with complex Morlet wavelet on O2 signal using the same setting of window in Section 4.2. Then, the threshold was compared with the curve to detect start and end points of alpha waves on O2 signal. Once an end point was detected, features from its corresponding VEOG signal was extracted and put into the trained LSTM classifier.

4.4. Classifying end points using LSTM

The purpose of the LSTM classifier is to determine whether the detected end point is the end point of alpha waves in alpha blocking phenomenon or in alpha wave attenuation-disappearance phenomenon. Thus, the key to distinguishing those two end points lies in differentiating the corresponding waveform on VEOG signal around the end point.

4.4.1. Feature extraction

For VEOG signal, we chose Haar wavelet as the mother wavelet for CWT, because its square-shaped functions are similar to VEOG signal during eye-closure events, as shown in Fig. 6(b). On CWT of VEOG in Fig. 4, the wavelet coefficients around e^1 are large (yellow bars), while those around p^2 are small (blue area). This is the basis for differentiating different end points. For feature extraction, Haar wavelet scaling from 1 to 128 was applied to the 0.5 s window before the end points p^2 and e^1, which was either [e^1 – 0.5 s, e^1 s] or [p^2 – 0.5 s, p^2 s]. And the corresponding wavelet energy values at each scale form an 128-dimensional wavelet energy

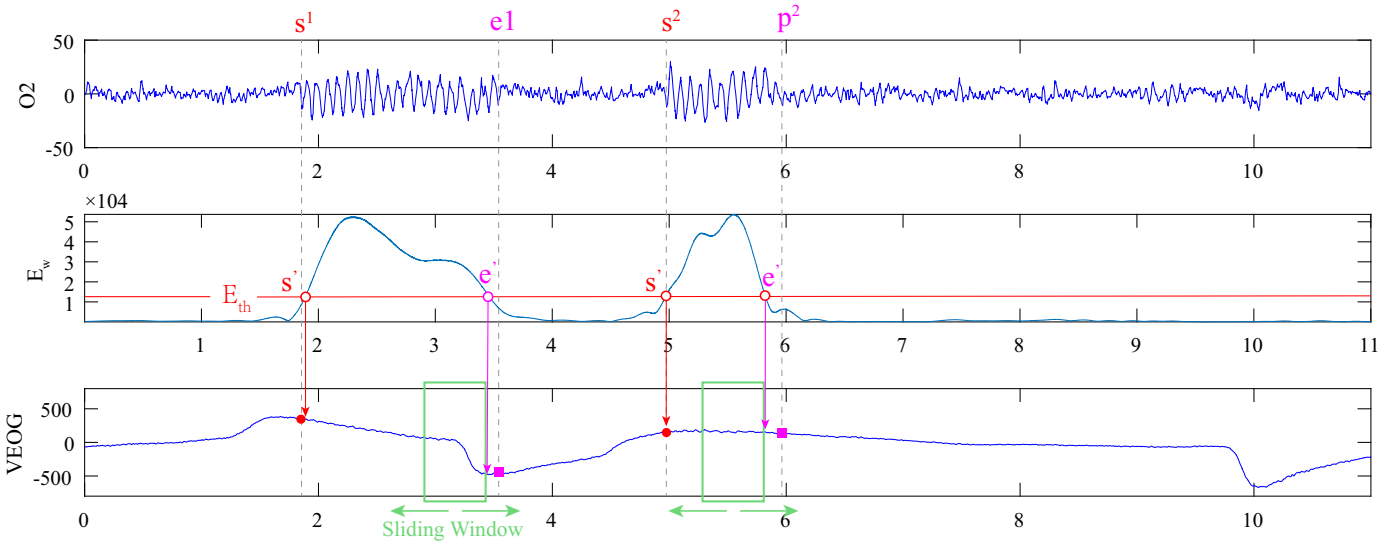


Fig. 8. Driver sleepiness detection based on O2 and VEOG signals. s^1 and s^2 are the actual start points of alpha waves, and e^1 and p^2 are the actual end points of alpha waves. s' and e' are the start and end points detected by alpha wavelet energy threshold. E_w is alpha wavelet energy curve.

feature vector. Moreover, to make full use of temporal information, the 0.5 s window was moved forward and backward for two times respectively, and the sliding step size was 0.125 s. After applying the same feature extraction method on each window, we obtained five sequential feature vectors from VEOG signal.

All the five feature vectors of the end point p^2 were marked as positive samples, and those of the end point e^1 were marked as negative samples. Depending on different training and test strategies, these feature vectors were either directly put into the trained LSTM classifier, or were augmented by CWGAN and then put into the classifier.

4.4.2. CWGAN for EOG data augmentation

Luo and Lu proposed CWGAN framework for EEG data augmentation to enhance EEG-based emotion recognition, and their models trained on appending datasets made significant improvements in classification accuracy [26]. Therefore, we adopt CWGAN to augment EOG samples for training LSTM classifier. CWGAN is used to generate artificial EOG data in the form of 128-dimensional feature vector from noise distribution.

Let X_r and X_g denote real and generated data distributions. The generator G produces X_g that resembles X_r in order to 'fool' the discriminator D, while the discriminator D discriminates between samples from real distribution X_g and those produced by generator G. The goal can then be formulated as a minmax problem:

$$\min_{\theta_G}\max_{\theta_D} L(X_r, X_g) = \mathbb{E}_{x_r\sim X_r}[\log(D(x_r))] + \mathbb{E}_{x_g\sim X_g}[\log(1 - D(x_g))] \quad (4)$$

where θ_G and θ_D denote parameters in generator and discriminator, respectively, $X_g = G(x_z)$, and x_z is sampled from a noise distribution.

The discontinuous property of Jenson-Shannon divergence in traditional adversarial training of GANs makes it difficult to formulate a gradient to optimize the generator. Therefore, CWGAN applied Kantorovich-Rubinstein duality of Earth-Mover distance (EMD, also called Wasserstein-1):

$$W(X_r, X_g) = \frac{1}{K}\sup_{||f||_L\le K}\mathbb{E}_{x_r\sim X_r}[f(x_r)] - \mathbb{E}_{x_g\sim X_g}[f(x_g)] \quad (5)$$

where f, which denotes the set of 1-Lipschitz functions, is replaced with discriminator D in realistic implementations, and $||f||_L \le K$ is replaced with $||D||_L \le 1$.

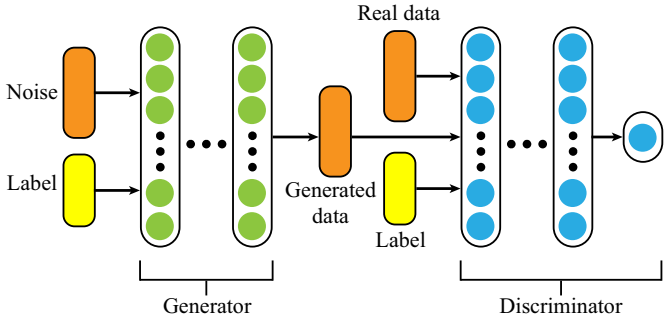


Fig. 9. Structure of CWGAN.

Meanwhile, instead of weight clipping, Lipschitz constraint with gradient penalty is used to directly constrain the gradient norm [51]. And, an auxiliary label Y_r is fed into the generator and discriminator, and it is concatenated with X_z in the generator, and X_r and X_g in the discriminator, respectively. So, CWGAN can be formulated as follows:

$$\begin{aligned}&\min_{\theta_G}\max_{\theta_D} L(X_r, X_g, Y_r)\\ &= \mathbb{E}_{x_r\sim X_r, y_r\sim Y_r}[D(x_r|y_r)] - \mathbb{E}_{x_g\sim X_g, y_r\sim Y_r}[D(x_g|y_r)]\\ &\quad - \lambda\mathbb{E}_{\hat{x}\sim\hat{X}, y_r\sim Y_r}[(||\nabla_{\hat{x}|y_r}D(\hat{x}|y_r)||_2 - 1)^2]\end{aligned} \quad (6)$$

where λ controls the trade-off between original objective and gradient penalty, and $\hat{x}$ denotes the data points sampled from the straight line between X_r and X_g. The losses of discriminator and generator are optimized in an alternating manner. The structure of CWGAN is shown in Fig. 9.

The generated EOG features are considered to be of high quality using discriminator loss, because it represents EMD between X_r and X_g when the network converges [26].

4.4.3. LSTM Classifier

LSTM network, as a special variant of recurrent neural network, is known to learn problems with long-term temporal dependencies [52]. LSTM has recurrent connections between hidden unites that allow information to pass from one time step of the network to the next. It extracts temporal information using LSTM layer and learns feature representation from the input sequence. In detail, the LSTM layer consists of the cell states propagated through time.

Table 1
ECE^1 and ECE^2 detection performance.

subject	$\#ECE^1$	start point F1(%)	end point F1(%)	$\#ECE^2$	start point F1(%)	end point F1(%)
1	34	95.5	98.6	24	98.0	98.0
2	28	98.2	94.7	35	100.0	100.0
3	22	95.5	93.0	30	94.9	94.9
4	20	88.4	93.0	27	90.6	90.6
5	26	98.1	92.3	24	95.8	95.8
6	28	91.2	92.6	26	92.3	92.3
7	28	98.2	98.2	27	94.5	94.5
8	27	94.5	94.5	28	94.9	94.9
9	23	95.7	93.3	26	98.0	98.0
10	21	95.0	97.6	23	97.9	97.9
11	25	96.0	96.2	29	98.2	98.2
12	26	96.3	94.1	28	94.5	94.5
Mean ± SD		95.2 ± 2.8	94.8 ± 2.2		94.6 ± 2.9	94.6 ± 2.9

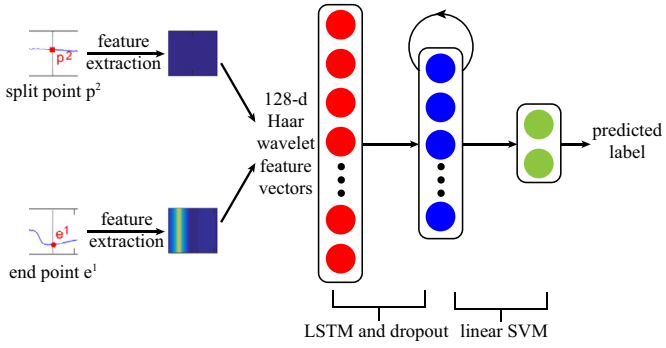


Fig. 10. Structure of the LSTM classifier.

At each time step, cell states are updated according to the gate output, which is calculated using the current input and the previous hidden states. The formulations are as follows:

$$\begin{aligned} f_t &= \sigma(W_f[h_{t-1}, x_t] + b_f) \\ i_t &= \sigma(W_i[h_{t-1}, x_t] + b_i) \\ g_t &= \tanh(W_g[h_{t-1}, x_t] + b_g) \\ c_t &= c_{t-1} * f_t + i_t * g_t, \end{aligned} \tag{7}$$

where σ denotes the sigmoid function, f_t and i_t denote the forget gate and input gate, respectively, g_t denotes the candidate of cell states, W_f, W_i and W_g are weight matrices, b_f, b_i and b_g are biases, and c_t and h_t are the cell states and the output of LSTM block. The forget gate controls the process of forgetting information by multiplying the cell states by real numbers between zero and one. Similarly, the input gate controls the process of remembering information. The current hidden states can then be calculated using the current cell states and the output of the output gate o_t. The formulations are as follows:

$$\begin{aligned} o_t &= \sigma(W_o[h_{t-1}, x_t] + b_o) \\ h_t &= o_t * \tanh(c_t). \end{aligned} \tag{8}$$

where W_o and b_o are the weight matrix and the bias for the output gate o_t, respectively.

For our LSTM classifier, dropout was adopted to avoid overfitting. Meanwhile, for the classification layer, Hinge loss with L2 regularization was taken as the loss function, so it can be considered as a linear kernel SVM. The structure of our classifier is shown in Fig. 10.

4.4.4. Judging end points of alpha waves

For the detected end point of alpha waves, we used the method mentioned in Section 4.4.1 to extract five 128-dimensional feature vectors from its corresponding VEOG signal. Then, the trained LSTM classifier produced 5 output labels corresponding to the five input feature vectors. And we chose the label that was the most frequent among all the output labels as the final label for the detected end point. This final label indicates whether the detected end point is the end point of alpha waves in alpha blocking phenomenon or in alpha wave attenuation-disappearance phenomenon.

5. Experiment and result

We first tested our model's performance in terms of how well it detects start and end points of alpha waves. Then, we compared different classifiers and feature extraction settings in order to better evaluate our proposed model. For feature extraction of VEOG signals, we tried different combinations of window size and sliding step size. For CWGAN, we tested its performance on the LSTM classifier, SVM and k-NN using subject-to-subject training and test, because samples are not sufficient under this scenario. And the CWGAN-based data augmentation method was also compared with another data augmentation approach similar to SMOTE [53]. To further evaluate the performance of our proposed LSTM classifier, we compared it with RNN, SVM and k-NN using leave-one-subject-out cross validation strategy.

5.1. Performance of detecting start and end points of alpha waves

As shown on E_w curve in Fig. 8, the detection of end points is sensitive to alpha wave energy threshold, resulting in the deviation of the detected end point from the actual end point. Therefore, we did not enforce exact detection of end points. Instead, if the detected start point s^1 and s^2 or end point e^1 fell into the range of $[s^1(s^2) - 0.5\ \text{s},\ s^1(s^2) + 0.5\ \text{s}]$ or $[e^1 - 0.5\ \text{s},\ e^1 + 0.5\ \text{s}]$, the point was regarded as a correctly detected point. If the detected end point p^2 of alpha waves fell into the range of $[p^2 - 0.8\ \text{s},\ p^2 + 0.8\ \text{s}]$, it was regarded as correctly detected point. We set a larger range for p^2, because the exact split point in alpha wave attenuation-disappearance was usually hard to determine.

We used F1 score to evaluate our model's performance on detecting start and end points of alpha waves. From Table 1, we can see that the F1 score for detecting start and end points of alpha waves in ECE^1 was around 95%. For ECE^2, its F1 score was also around 95%. This demonstrates our model's strong ability in detecting start and end points of alpha waves.

5.2. Different combinations of window size and sliding step

We compared the performance of our LSTM classifier in terms of different combinations of window size and sliding step size.

Table 2
Accuracy(%) and standard deviation(%) of classifying end points of alpha waves with different combinations of window size and sliding step.

Subject	Window size = 0.2 Step size = 0.2	Window size = 0.4 Step size = 0.15	Window size = 0.5 Step size = 0.125	Window size = 0.6 Step size = 0.1
1	92.34	96.60	97.87	97.45
2	99.59	98.37	99.18	99.59
3	97.70	97.05	98.03	97.38
4	90.69	95.17	96.90	98.97
5	93.73	97.65	98.43	98.43
6	97.14	98.37	98.78	97.14
7	98.60	97.21	98.14	98.14
8	93.33	95.93	96.67	97.41
9	96.54	97.69	98.46	98.85
10	98.40	96.40	98.80	97.20
11	95.71	98.10	98.57	99.05
12	96.96	97.39	97.83	97.83
Averaged accuracy	95.90	97.16	**98.14**	98.12
SD	2.77	0.99	**0.75**	0.84

Leave-one-subject-out cross validation was used as the evaluation strategy. As listed in Table 2, among the settings we tried, there was no overlapping of sliding windows in the setting of 0.2 s window size with 0.2 s sliding step, while there was overlapping in all the other combinations. Meanwhile, the total coverage of all windows over time was set to 1 s for each setting so that the amount of information contained in all the settings was the same.

From Table 2, we can see that all the windows with overlapping outperformed the window without overlapping. This indicates that the context included in the feature vector can improve the performance of classification. And, 0.5 s window with 0.125 s sliding step size was the best. Therefore, we chose it to be our window setting for feature extraction of VEOG signal.

5.3. Details of data separation, data augmentation and classifiers

5.3.1. Leave-one-subject-out

For each subject, we used VEOG signal from simulated driving experiments and repeated eye closure experiments. In simulated driving experiments, there were usually very few eye closure events in the early stage. In the late stage, however, the two eye closure events occurred more often as the subject became more sleepy. Therefore, we picked a 30 min period with frequent appearance of the two alpha-related phenomena. In this period, the amount of end points e^1 and p^2 were almost balanced. Besides, to augment training samples, we sequentially picked some of the end points e^1 in repeated eye closure experiments. And we ensured that the training set was not significantly unbalanced after we added those end points. For each subject, there were about 80 end points in the dataset, among which the number of p^2 was around 30.

Then, we left one subject's data out, and took the remaining 11 subjects' data to train the LSTM classifier. The classifier was then tested on the one subject's data. We did this training and test process in turns, and used the averaged result to evaluate the performance of our classifier.

5.3.2. Subject-to-subject evaluation

For training and test on each subject, we calculated the number of samples we picked for each subject in Section 5.3.1, and decreased them to the same amount as the subject with minimum sample number. However, the number of samples was not sufficient for training the LSTM classifier, so we used two different data augmentation methods. One is an oversampling method similar to SMOTE, and the other is CWGAN [53].

For the first method, the way we augmented the dataset is similar to manually enforce a deviation on end points, which also allows us to evaluate the classifier's performance when an end point is deviated from the actual one. In the original dataset, for an end point at time point t_e, we picked N points evenly distributed in range $[t_e - 0.1\text{ s}, t_e\text{ s})$ and $(t_e\text{ s}, t_e + 0.1\text{ s}]$, respectively, into the dataset, as shown in Fig. 11. This range was chosen according to the duration of downward trend line around end point e^1. Together with the original end point, the augmented dataset has $2N + 1$ samples. Then, we split the augmented dataset into training and test sets, where the test set only included samples from the simulated driving experiment, because we would like to test our model under a simulated driving condition.

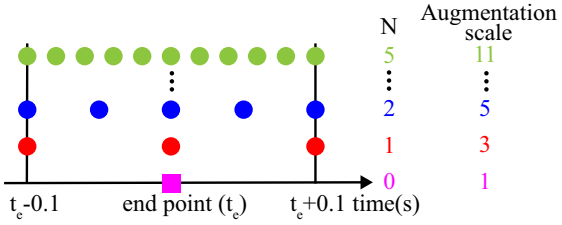


Fig. 11. Illustration of SMOTE-like augmentation method. N is the number of points evenly distributed in time range $[t_e - 0.1\text{ s}, t_e\text{ s})$, equivalent to those in time range $(t_e\text{ s}, t_e + 0.1\text{ s}]$. Augmentation scale corresponds to the Scale column in Table 4.

For the second method, CWGAN, we first separated the original data for each subject into training and test sets (denote the number of samples in the training set as N_1) for k-NN, SVM and LSTM classifier. Then, the training data was further split into training and testing set (denote the number of samples as N_2 and N_3, respectively, and $N_1 = N_2 + N_3$) for CWGAN. We evaluated the performance of CWGAN with different numbers of data appended to original dataset.

5.3.3. Evaluation details for CWGAN

ReLU activation function was used for all hidden layers, and Adam optimizer was chosen to train the networks. λ was set to 10, and the noise followed a uniform distribution $U[-1, 1]$. Meanwhile, SVM was used as the classifier for CWGAN, and the parameter C was tuned to find the optimal value. When training CWGAN, we observed that the loss of discriminator always converged quickly, indicating that the generated data are of high quality.

5.3.4. Evaluation details for the classifiers

For our proposed LSTM classifier, Adam optimizer was chosen to train the network. And we selected dozens of hyper-parameters from a given range. Table 3 listed the hyper-parameters in our

Table 3
The hyper-parameters and their range of the LSTM classifier.

hyper-parameter	range
hidden size	16 ~ 128
dropout probability	0.2 ~ 0.9
$\log_{10}$(L2 regularization strength)	-7 ~ -2
$\log_{10}$(learning rate)	-5 ~ -1

Table 4
Accuracy(%) of *k*-NN, SVM and LSTM for classifying end points of alpha waves with SMOTE-like augmentation method. Scale denotes the augmentation scale of the dataset.

Scale	*k*-NN	SVM	LSTM
1	**91.58**	**95.03**	**93.98**
3	66.23	80.78	88.50
5	61.28	71.88	88.18
7	60.38	69.17	88.33
9	59.74	67.03	88.88
11	59.11	65.26	87.61

model and their ranges. We adopted 10-fold cross validation to choose the best sets of hyper-parameters.

The settings for RNN were same as LSTM. We implemented SVM and *k*-NN by sklearn module in Python, and the penalty parameter *C* in SVM and *k* in *k*-NN were tuned. *k* was set to 5.

5.4. Comparison for two data augmentation methods

As shown in Table 4, for *k*-NN, SVM and LSTM classifier, the SMOTE method cannot improve the classification accuracy. With more number of augmented data, the accuracy tends to decline in general. This decreasing trend was normal, because there are more deviated end points with the growth of sample numbers. *k*-NN and SVM suffer from a much larger decrease than LSTM, indicating that LSTM can alleviate the impact of deviation on classification accuracy. Meanwhile, the accuracy drops drastically when the size of the augmented dataset is 3 or 5 times the size of the original dataset. In contrary, the accuracy becomes stable when the size of the augmented dataset continues to grow. One reason may be that the end points become closer to each other when the size of the dataset is augmented to a certain degree, i.e. 5 in our model.

The accuracies of *k*-NN, SVM and LSTM classifier for CWGAN augmentation method is shown in Table 4. For *k*-NN, CWGAN can increase the accuracy, and it reaches the highest accuracy with augmentation scale of 8. For SVM, there is little improvement in accuracy. When the augmented dataset is four or five times the size of the original dataset, SVM reached its highest accuracy with only 0.01% improvement. This may indicate that SVM can characterize distribution of the dataset with few samples. Therefore, when the generated data including noise are added, it does not help with the improvement in performance. For the LSTM classifier, we can observe improvement in all different sizes of appended data, and it has the highest accuracy of 96.61% with six times augmentation. Compared with *k*-NN and SVM, LSTM as a deep learning approach, need large-scale data to depict the true distribution of dataset. Thus, with much more augmented data, there are more boost in classification accuracy.

5.5. Comparison of classifiers

To make the comparison of classifiers simpler, we only took into account those correctly detected points in Section 5.1.

5.5.1. Subject-to-subject evaluation

We compared our LSTM classifier with *k*-NN and SVM using subject-to-subject evaluation. As shown in Table 5, with CWGAN data augmentation, our LSTM classifier has a much better accuracy of 96.61%. As mentioned in Section 5.1, the detection of end points is sensitive to alpha wave energy threshold, which results in deviation of the detected end point from the actual end point. Meanwhile, for SMOTE-like data augmentation, the way we augmented the dataset resembles the deviation manner of end points. And LSTM suffers from much less decrease in classification accuracy when the dataset size grows, as shown in Table 4. Therefore, LSTM can reduce the effect of deviation on classification performance by utilizing temporal information from a sequential input.

Table 5
Accuracy(%) of *k*-NN, SVM and LSTM for classifying end points of alpha waves with CWGAN augmentation method. Scale denotes the augmentation scale of the dataset.

Scale	*k*-NN	SVM	LSTM
1	91.58	95.03	93.98
2	94.21	95.03	95.43
3	93.84	94.77	96.09
4	92.26	**95.04**	94.66
5	93.96	**95.04**	95.32
6	93.96	94.91	**96.61**
7	93.57	94.90	94.90
8	**94.25**	94.91	94.93
9	93.18	94.64	94.79

Table 6
Accuracies and standard deviations of different classifiers under leave-one-subject-out cross validation.

Subject	k-NN	SVM	RNN	LSTM
1	90.34	93.44	98.04	96.90
2	86.56	95.74	98.04	98.03
3	86.53	95.92	97.48	98.78
4	90.70	97.21	98.63	98.14
5	85.11	95.74	99.21	97.87
6	87.76	96.33	99.20	99.18
7	89.80	92.16	98.45	98.43
8	90.74	92.96	98.52	96.67
9	93.48	95.22	95.77	97.83
10	92.86	95.71	97.20	98.57
11	88.85	95.77	98.10	98.46
12	89.60	94.80	98.70	98.80
Average accuracy	89.35	95.08	98.11	**98.14**
SD	2.54	1.49	0.95	**0.75**

5.5.2. Leave-one-subject-out cross validation

As shown in Table 6, our LSTM classifier achieves the highest accuracy of 98.14% and the lowest standard deviation of 0.75%, showing its strong classification performance and robustness. The overall performance of RNN is very close to that of the LSTM classifier, and its accuracy is slightly less than LSTM. Besides, *k*-NN model in which *k* was 5 is the worst among them, because it just simply classifies a point according to its neighbors without considering temporal information. Meanwhile, although the averaged accuracy of SVM is close the LSTM, it is less stable across subjects in terms of standard deviation. We ascribe the robustness of our LSTM classifier to its recurrent structure that makes use of sequential information. Therefore, even if the information carried in some of the five input feature vectors is incomplete, our LSTM classifier can still correctly classify the end points by integrating temporal information. What's more, LSTM can deal with the deviation problem mentioned in Section 5.5.1. If the detected point slightly deviates from the true end point, it can still fall into the range of the sliding window. All these advantages of LSTM leads to its satisfying performance.

6. Conclusion and future work

In this paper, we have proposed a novel method for driver sleepiness detection from EEG and EOG signals. Our model aims to track the change of alpha waves on O2 signal and differentiate the two alpha-related phenomena. The proposed model utilizes continuous wavelet transform to extract features from physiological signals, CWGAN to augment EOG features in dataset, and LSTM to classify end points of alpha waves. Our results have demonstrated that our model can detect start and end points of alpha waves with high accuracy. CWGAN can effectively generate realistic-like EOG data and increase classification accuracy. Meanwhile, the LSTM classifier outperforms SVM and *k*-NN classifiers in both subject-to-subject evaluation and leave-one-subject-out cross validation. As our model places very few electrodes on the subject, we believe that it is practical for routine use in real-life scenarios.

Our future work will focus on facilitating the feature extraction process. If continuous wavelet transform can be replaced with deep neural networks such as CNN, we no longer need to extract features manually. Meanwhile, we will adopt transfer learning methods to deal with individual differences across subjects [54–56].

Declaration of Competing Interest

None.

Acknowledgments

This work was supported in part by grants from the National Key Research and Development Program of China (Grant no. 2017YFB1002501), the National Natural Science Foundation of China (Grant no. 61673266), the Technology Research and Development Program of China Railway Corporation (Grant No.2016Z003-B), and the Fundamental Research Funds for the Central Universities.

References

[1] H. Hakkanen, H. Summala, Sleepiness at work among commercial truck drivers, Sleep 23 (1) (2000) 49–57.
[2] D.F. Dinges, An overview of sleepiness and accidents., J. Sleep Res. 4 (1995) 4–14.
[3] F. Sagberg, Road accidents caused by drivers falling asleep, Accid. Anal. Prev. 31 (6) (1999) 639–649.
[4] A. Sahayadhas, K. Sundaraj, M. Murugappan, Detecting driver drowsiness based on sensors: a review, Sensors 12 (12) (2012) 16937–16953.
[5] R.D. Ogilvie, The process of falling asleep, Sleep Med. Rev. 5 (3) (2001) 247–270.
[6] S.K.L. Lal, A. Craig, A critical review of the psychophysiology of driver fatigue, Biol. Psychol. 55 (3) (2001) 173–194.
[7] J.L. Cantero, M. Atienza, R.M. Salas, Human alpha oscillations in wakefulness, drowsiness period, and rem sleep: different electroencephalographic phenomena within the alpha band, Neurophysiol. Clin. Clin. Neurophysiol. 32 (1) (2002) 54–71.
[8] Y. Dong, Z. Hu, K. Uchimura, N. Murayama, Driver inattention monitoring system for intelligent vehicles: a review, IEEE Trans. Intell. Transp. Syst. 12 (2) (2011) 596–614.
[9] R.N. Khushaba, S. Kodagoda, S. Lal, G. Dissanayake, Driver drowsiness classification using fuzzy wavelet-packet-based feature-extraction algorithm, IEEE Trans. Biomed. Eng. 58 (1) (2011) 121–131.
[10] R.P. Balandong, R.F. Ahmad, M.N.M. Saad, A.S. Malik, A review on eeg-based automatic sleepiness detection systems for driver, IEEE Access 6 (2018) 22908–22919.
[11] C.C. Liu, S. Hosking, M.G. Lenne, Predicting driver drowsiness using vehicle measures: recent insights and future challenges., J. Saf. Res. 40 (4) (2009) 239–245.
[12] A.C. Guyton, Structure and Function of the Nervous System, Saunders Limited., 1976.
[13] P. Lo, M. Huang, K. Chang, Eeg alpha blocking correlated with perception of inner light during zen meditation, Am. J. Chin. Med. 31 (4) (2003) 629–642.
[14] C.J. Harland, T.D. Clark, R.J. Prance, Remote detection of human electroencephalograms using ultrahigh input impedance electric potential sensors, Appl. Phys. Lett. 81 (17) (2002) 3284–3286.
[15] Y. Jiao, B.-L. Lu, An alpha wave pattern from attenuation to disappearance for predicting the entry into sleep during simulated driving, in: Proceedings of the Eighth International IEEE/EMBS Conference on Neural Engineering (NER), 2017, pp. 21–24, doi:10.1109/NER.2017.8008282.
[16] Y. Jiao, B.-L. Lu, Detecting driver sleepiness from eeg alpha wave during daytime driving, in: Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2017, pp. 728–731, doi:10.1109/BIBM.2017.8217744.
[17] M.H. Silber, S. Ancoliisrael, M.H. Bonnet, S. Chokroverty, M.M. Griggdamberger, M. Hirshkowitz, S. Kapen, S. Keenan, M.H. Kryger, T. Penzel, et al., The visual scoring of sleep in adults., J. Clin. Sleep Med. 3 (2) (2007) 121–131.
[18] J. Liu, H. Ding, A. Shahroudy, L. Duan, X. Jiang, G. Wang, A. Kot Chichung, Feature boosting network for 3D pose estimation, IEEE Trans. Pattern Anal. Mach. Intell. (2019) 1, doi:10.1109/TPAMI.2019.2894422.
[19] J. Liu, A. Shahroudy, D. Xu, A.C. Kot, G. Wang, Skeleton-based action recognition using spatio-temporal LSTM network with trust gates, IEEE Trans. Pattern Anal. Mach. Intell. 40 (12) (2018) 3007–3021, doi:10.1109/TPAMI.2017.2771306.
[20] J. Liu, G. Wang, P. Hu, L.-Y. Duan, A.C. Kot, Global Context-aware Attention LSTM Networks for 3D Action Recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
[21] H. Li, J. Liu, G. Zhang, Y. Gao, Y. Wu, Multi-glimpse LSTM with Color-depth Feature Fusion for Human Detection, in: Proceedings of the IEEE International Conference on Image Processing (ICIP), 2017, pp. 905–909, doi:10.1109/ICIP.2017.8296412.
[22] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436.
[23] S. Min, B. Lee, S. Yoon, Deep learning in bioinformatics, Brief. Bioinform. 18 (5) (2017) 851–869.
[24] H. Dong, A. Supratak, W. Pan, C. Wu, P.M. Matthews, Y. Guo, Mixed neural network approach for temporal sleep stage classification, IEEE Trans. Neural Syst. Rehabil. Eng. 26 (2) (2018) 324–333.
[25] A. Supratak, H. Dong, C. Wu, Y. Guo, Deepsleepnet: a model for automatic sleep stage scoring based on raw single-channel eeg, IEEE Trans. Neural Syst. Rehabil. Eng. 25 (11) (2017) 1998–2008.
[26] Y. Luo, B.-L. Lu, Eeg data augmentation for emotion recognition using a conditional Wasserstein Gan, in: Proceedings of the Fortieth Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2018, pp. 2535–2538.
[27] A. Radford, L. Metz, S. Chintala, Unsupervised Representation Learning With Deep Convolutional Generative Adversarial Networks, arXiv:1511.06434 (2015).
[28] C. Lin, C. Chuang, C. Huang, S.W. Tsai, S. Lu, Y. Chen, L. Ko, Wireless and wearable eeg system for evaluating driver vigilance, IEEE Trans. Biomed. Circuits Syst. 8 (2) (2014) 165–176.
[29] H.J. Eoh, M.K. Chung, S. Kim, Electroencephalographic study of drowsiness in simulated driving with sleep deprivation, Int. J. Ind. Ergon. 35 (4) (2005) 307–320.
[30] G. Li, B. Lee, W. Chung, Smartwatch-based wearable eeg system for driver drowsiness detection, IEEE Sens. J. 15 (12) (2015) 7169–7180.
[31] L.-C. Shi, B.-L. Lu, Eeg-based vigilance estimation using extreme learning machines, Neurocomputing 102 (2013) 135–143.
[32] D. Wu, V.J. Lawhern, S.M. Gordon, B.J. Lance, C. Lin, Driver drowsiness estimation from eeg signals using online weighted adaptation regularization for regression (owarr), IEEE Trans. Fuzzy Syst. 25 (6) (2017) 1522–1535.
[33] C.-T. Lin, K.-C. Huang, C.-H. Chuang, L.-W. Ko, T.-P. Jung, Can arousing feedback rectify lapses in driving? prediction from eeg power spectra, J. Neural Eng. 10 (5) (2013) 056024.
[34] B. Jammes, H. Sharabty, D. Esteve, Automatic eog analysis: a first step toward automatic drowsiness scoring during wake- sleep transitions, Somnologie Schlafforschung Und Schlafmedizin 12 (3) (2008) 227–232.
[35] J.-X. Ma, L.-C. Shi, B.-L. Lu, Vigilance estimation by using electrooculographic features, in: Proceedings of the Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE, IEEE, 2010, pp. 6591–6594.
[36] Y. Zhang, X. Gao, J. Zhu, W. Zheng, B.-L. Lu, A novel approach to driving fatigue detection using forehead EOG, in: Proceedings of the Seventh International IEEE/EMBS Conference on Neural Engineering (NER), 2015, pp. 707–710, doi:10.1109/NER.2015.7146721.
[37] W.-L. Zheng, B.-L. Lu, A multimodal approach to estimating vigilance using EEG and forehead EOG, J. Neural Eng. 14 (2) (2017) 026017.
[38] V. Kartsch, S. Benatti, D. Rossi, L. Benini, A wearable eeg-based drowsiness detection system with blink duration and alpha waves analysis, in: Proceedings of the Eighth International IEEE/EMBS Conference on Neural Engineering (NER), 2017, pp. 251–254, doi:10.1109/NER.2017.8008338.
[39] L. Chen, Y. Zhao, J. Zhang, J. Zou, Automatic detection of alertness/drowsiness from physiological signals using wavelet-based nonlinear features and machine learning, Expert Syst. Appl. 42 (21) (2015) 7344–7355.
[40] J. Arnin, D. Anopas, M. Horapong, P. Triponyuwasi, T. Yamsa-ard, S. Iampetch, Y. Wongsawat, Wireless-based portable EEG-EOG monitoring for real time drowsiness detection, in: Proceedings of the Thirty-Fifth Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2013, pp. 4977–4980, doi:10.1109/EMBC.2013.6610665.
[41] W. Zheng, K. Gao, G. Li, W. Liu, C. Liu, J. Liu, G. Wang, B.-L. Lu, Vigilance estimation using a wearable EOG device in real driving environment, IEEE Trans. Intell. Transp. Syst. (2019) 1–15,, doi:10.1109/TITS.2018.2889962.

[42] H. Tang, W. Liu, W.-L. Zheng, B.-L. Lu, Multimodal emotion recognition using deep neural networks, in: D. Liu, S. Xie, Y. Li, D. Zhao, E.-S. M. El-Alfy (Eds.), Neural Information Processing, Springer International Publishing, Cham, 2017, pp. 811–819.
[43] M.W. Johns, A new method for measuring daytime sleepiness: the epworth sleepiness scale., Sleep 14 (6) (1991) 540–545.
[44] P.A. Kanda, E.F. Oliveira, F.J. Fraga, Eeg epochs with less alpha rhythm improve discrimination of mild alzheimer's, Comput. Methods Programs Biomed. 138 (2017) 13–22.
[45] S. Hanslmayr, J. Gross, W. Klimesch, K.L. Shapiro, The role of alpha oscillations in temporal attention, Brain Res. Rev. 67 (1–2) (2011) 331–343.
[46] Y. Deng, Y. Jiao, B.-L. Lu, Driver sleepiness detection using LSTM neural network, in: Proceedings of the International Conference on Neural Information Processing, Springer, 2018, pp. 622–633.
[47] R. Naga, S. Chandralingam, T. Anjaneyulu, K. Satyanarayana, Denoising eog signal using stationary wavelet transform, Meas. Sci. Rev. 12 (2) (2012) 46–51.
[48] M.S. Reddy, B. Narasimha, E. Suresh, K.S. Rao, Analysis of eog signals using wavelet transform for detecting eye blinks, in: Proceedings of the International Conference on Wireless Communications Signal Processing (WCSP), 2010, pp. 1–4, doi:10.1109/WCSP.2010.5633797.
[49] A.N. Pavlov, A.E. Hramov, A.A. Koronovskii, E.Y.E. Sitnikova, V.A. Makarov, A.A. Ovchinnikov, Wavelet analysis in neurodynamics, Phys. Usp. 55 (9) (2012) 845–875.
[50] G. Li, B. Lee, W. Chung, Smartwatch-based wearable eeg system for driver drowsiness detection, IEEE Sens. J. 15 (12) (2015) 7169–7180.
[51] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, A.C. Courville, Improved training of Wasserstein Gans, in: Proceedings of the Advances in Neural Information Processing Systems, 2017, pp. 5767–5777.
[52] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (8) (1997) 1735–1780.
[53] N.V. Chawla, K.W. Bowyer, L.O. Hall, W.P. Kegelmeyer, Smote: synthetic minority over-sampling technique, J. Artif. Intell. Res. 16 (1) (2002) 321–357.
[54] H. Li, W.-L. Zheng, B.-L. Lu, Multimodal vigilance estimation with adversarial domain adaptation networks, in: Proceedings of the International Joint Conference on Neural Networks (IJCNN), IEEE, 2018, pp. 1–6.
[55] H. Li, Y.-M. Jin, W.-L. Zheng, B.-L. Lu, Cross-subject emotion recognition using deep adaptation networks, in: Proceedings of the International Conference on Neural Information Processing, Springer, 2018, pp. 403–413.
[56] Y. Lu, W.-L. Zheng, B. Li, B.-L. Lu, Combining eye movements and eeg to enhance emotion recognition., in: Proceedings of the IJCAI, 15, 2015, pp. 1170–1176.

Yingying Jiao received the M.S. degree in computer science with the Department of Computer Science and Technology, Anhui University, Hefei, China. She is currently pursuing the Ph.D degree in computer science with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China. Her current research interests include driver fatigue detection, brain-computer interaction, machining learning, and pattern recognition.

Yini Deng received the B.S. degree in computer science and technology from Shanghai Jiao Tong University, Shanghai, China. She is currently pursuing the master degree in computer science in Shanghai Jiao Tong University. Her current research interests include driver sleepiness detection, neural networks, and machine learning.

Yun Luo received the B.S. degree in software engineering from University of Electronic Science and Technology of China, Chengdu, China. He is currently pursuing the Ph.D degree in computer science with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China. His current research interests include driver fatigue detection, emotion recognition, sleep quality evaluation, transfer learning, and deep neural networks.

Bao-Liang Lu received the B.S. degree in instrument and control engineering from the Qingdao University of Science and Technology, Qingdao, China, in 1982, the M.S. degree in computer science and technology from Northwestern Polytechnical University, Xi'an, China, in 1989, and the Dr. Eng. degree in electrical engineering from Kyoto University, Kyoto, Japan, in 1994. He was with the Qingdao University of Science and Technology from 1982 to 1986. From 1994 to 1999, he was a Frontier Researcher with the Bio-Mimetic Control Research Center, Institute of Physical and Chemical Research (RIKEN), Nagoya, Japan, and a Research Scientist with the RIKEN Brain Science Institute, Wako, Japan, from 1999 to 2002. Since 2002, he has been a Full Professor with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China. He received the IEEE TRANSACTIONS on AUTONOMOUS MENTAL DEVELOPMENT Outstanding Paper Award from the IEEE Computational Intelligence Society in 2018. His current research interests include brain-like computing, neural networks, machine learning, brain computer interaction, and affective computing. Prof. Lu is currently a Board Member of the Asia Pacific Neural Network Society (APNNS, previously APNNA) and the Steering Committee member of the IEEE TRANSACTIONS ON AFFECTIVE COMPUTING. He was the President of the Asia Pacific Neural Network Assembly (APNNA) and the General Chair of the 18th International Conference on Neural Information Processing in 2011. He is currently the Associate Editor of the IEEE TRANSACTIONS ON COGNITIVE AND DEVELOPMENT SYSTEMS.